THE UNIVERSITY *of* EDINBURGH

High-Resolution Structural Studies of Kynurenine 3-Monooxygenase

Mark R.D. Taylor

Supervised by Dr Chris Mowat

Doctor of Philosophy

The University of Edinburgh

2017

Acknowledgements

I would like to thank Dr Chris Mowat for his guidance throughout the years and making this all possible. Thanks to Graham Moran, University of Minnesota, for providing the *Pf*KMO gene. Thanks to Dr Helen Bell and Dr Martin Wilkinson for their previous work on dm2 and providing the kinetic data presented throughout this thesis. Thanks to Dr Atlanta Cook and Dr Julia Richardson for helping me with synchrotron data collection. Thanks to Dr. Logan MacKay and Dr. Faye Cruickshank for their assistance with collecting mass spectroscopy data. I would also like to thank all others who gave their time and assistance throughout my PhD including Dr Helen Bell, the entire Campopiano group, in particular Dr Chris McLean, all my Masters students and the Cockroft group.

Declaration

I declare that the work presented in this thesis is my own except where clearly indicated. This work has not been submitted for any other degree.

Lay Summary

95% of the essential amino acid L-tryptophan (L-Trp) enters the kynurenine pathway (KP). The molecules produced in this pathway are known as the kynurenines and are associated with a number of diseases including Alzheimer's disease, Huntington's disease, cancer and acute pancreatitis. The enzyme kynurenine 3-monooxygenase (KMO) is part of the KP and catalyses the formation of 3-hydroxy-L-kynurenine (3HK) from L-kynurenine (L-Kyn). Studies on animal models have shown the inhibition of KMO activity alleviates symptoms for a number of diseases associated with the KP. Due to difficulty in isolating human KMO much research has utilised KMO from the bacterium *Pseudomonas fluorescens* (*Pf*KMO) as a model. Greater understanding of this protein could give rise to more effective treatments.

This thesis describes the study of a modified *Pf*KMO enzyme, called dm2. Dm2 is a more soluble protein than *Pf*KMO and its behaviour is similar enough to make it an adequate model. The work is based upon x-ray crystallography, creating 3D models of dm2 and its variants. These 3D models allow us to view the interactions between dm2 and the molecule bound within it. The role(s) of different sections of dm2 can also be studied. This can include the building blocks of dm2, the amino acids, and the larger sections of the protein, the domains.

Within this work amino acids within dm2's active site, concerned with binding L-Kyn, were altered. The 3D models and catalytic efficiencies of these altered proteins were compared to dm2 and the importance of specific amino acids was confirmed.

In previous work it was recognised that movement between two domains of dm2 gave rise to *closed* and *open* forms. It was hypothesised that the presence of L-Kyn caused dm2 to take on the *closed* form. This could have large implications in terms of the function of dm2. The 3D models with L-Kyn bound, however, came from crystals grown in a different form. This allowed for the possibility that the different crystal forms caused the movement of domains, rather than the presence of L-Kyn.

The crystals grown in this work were all of the same form regardless of L-Kyn exposure and are therefore comparable. In L-Kyn-exposed and L-Kyn-free models two molecules are seen; one *open*, the other *closed*. This suggests that L-Kyn is not responsible for the movement of domains. However L-Kyn is only seen in the protein when in the *closed* form. The *open* molecule contains no L-Kyn. This therefore does not diminish the importance of the movement of the domains.

In an attempt to study this movement amino acids were altered to inhibit dm2's ability to take on the *closed* form. The results of this were inconclusive since the *closed* form was still present within the model. The levels of L-Kyn present were, however, severely affected, even in the *closed* form. This shows the role of the domain movement is not fully understood and requires further study.

Abstract

The kynurenine pathway produces NAD^+ from L-tryptophan. Metabolites known as the kynurenines are produced within the pathway. The effects of the kynurenines have been associated with a number of diseases including cancer, Alzheimer's disease, Huntington's disease, and acute pancreatitis. Kynurenine monooxygenase (KMO) is an enzyme that catalyses the conversion of L-kynurenine to 3-hydroxy-L-kynurenine, the downstream product of which is the neurotoxic quinolinic acid.

L-kynurenine is positioned at a branching point within the pathway. Metabolism via KMO leads to quinolinic acid production whereas conversion via kynurenine aminotransferase (KAT) produces the neuroprotective kynurenic acid. Inhibition of KMO leads to an increase in kynurenic acid concentration. This has also been shown to ameliorate the symptoms of neurological diseases in a number of animal models as well as to protect against multiple organ dysfunction caused by acute pancreatitis in rodent models. These findings present KMO as a promising drug target.

Due to the hydrophobic nature of human KMO (hKMO) it has been necessary to utilise other forms of KMO as models. Past studies have produced crystal structures of a truncated *Saccharomyces cerevisiae* KMO and of *Pseudomonas fluorescens* KMO (*Pf*KMO). Previous work in this research group has resulted in the structure of variants of *Pf*KMO bound to either inhibitor molecules or substrate. These structures identified residues involved in substrate binding and the presence of a highly mobile section of the C-terminus, giving rise to *open* and *closed* conformations. It was surmised the movement of the C-terminus was dependent upon the presence of substrate and an interactive network between the C-terminus and the rest of the protein.

Using improved crystallising conditions high-resolution structures of *Pf*KMO have been produced that allow for further study of residues involved in substrate binding and the interactive network within the C-terminus.

The mutants R84K and Y404F showed severely decreased enzyme activity. Crystal structures of these proteins showed disrupted interactions between substrate and active site. These findings underline the importance of residues R84 and Y404 in substrate binding. An H320F mutation gives an analogous active site to hKMO. Crystallographic and kinetic study of this mutant proved very similar to *Pf*KMO, supporting the use of *Pf*KMO as a model for hKMO.

Throughout the work each structure had a $P2_122_1$ space group with two molecules in the asymmetric unit. The presence of an *open* and *closed* molecule within each structure, including substrate-free molecules refuted the connection between C-terminus and substrate. R386K and E372T mutations were separately introduced in order to interrupt the interactive network. The presence of both *open* and *closed* conformations in the structures of R386K and E372T refutes the necessity for the interactive network

in C-terminus movement. The data analysed throughout the project suggest simple mobility and thermal motion as the cause of the movement of the C-terminus.

This work, in conjunction with kinetic data from the thesis of Helen Bell, presents structural data to characterise the role of binding residues within the active site of KMO as well as the mechanistic role of the C-terminus. It also highlights the importance of certain binding residues and countered the previously held hypotheses surrounding the significance of the C-terminus. The mechanistic role of the C-terminus therefore remains unclear and requires further study.

Abbreviations

Amino acids:

Amino acid	Three letter code	One letter code
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamic acid	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

Nucleotides:

Nucleotide	One letter code
Adenine	A
Cytosine	C
Guanine	G
Thymine	T

General abbreviations:

1-DMT	1-methyl-D-tryptophan
1-LMT	1-methyl-L-tryptophan
1-MT	1-methyl-tryptophan
3-HA	3-hydroxyanthranilic acid
3-HAO	3-hydroxyanthranilic acid oxygenase
3-HK	3-hydroxy-L-kynurenine
°	Degrees
°C	Degrees Celsius
Å	Ångström
AD	Alzheimer's disease
AHR	Aryl hydrocarbon receptor
AMPAr	α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid receptor
AP	Acute pancreatitis
ASU	Asymmetric unit
Cα	Alpha carbon
$CC_{1/2}$	Correlation coefficient
CNS	Central nervous system
cyt b_5	cytochrome b_5
Da	Daltons
DC	Dendritic cell
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTPs	Deoxyribonucleotides
DTT	Dithiothreitol
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
FAD	Flavin adenine dinucleotide
FPLC	Fast protein liquid chromatography
g	Gravity
g (ng/µg/mg/kg)	Grams (nano/micro/milli/kilo)
GPR35	G protein-coupled receptor 35
h	Human
HD	Huntington's disease
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
IC_{50}	Half maximal inhibitory concentration

IDO1	Indoleamine 2,3-dioxygenase 1
IDO2	Indoleamine 2,3-dioxygenase 2
IFN-γ	Interferon gamma
IPTG	Isopropyl β-D-1-thiogalactopyranoside
KA	Kynurenic acid
KAT	Kynurenine aminotransferase
k_{cat}	Turnover number
K_i	Dissociation constant of inhibitor
K_M	Michaelis-Menten constant
KMO	Kynurenine 3-monooxygenase
KP	Kynurenine pathway
L (nL/µL/mL)	Litre (nano/micro/milli)
LB	Luria Bertani
L-Kyn	L-kynurenine
L-Trp	L-tryptophan
M (nM/µM/mM)	Molar (nano/micro/milli)
m (nm/mm)	Metre (nano/milli)
MB	Methylene blue
MHPCO	2-methyl-3-hydroxypyridine-5-carboxylic acid oxygenase
min	Minutes
MW	Molecular weight
nAChR	α7 nicotinic acetylcholine receptor
NAD/NAD^+	Nicotinamide adenine dinucleotide
NADPH/$NADP^+$	Nicotinamide adenine dinucleotide phosphate
NCS	Non-crystallography symmetry
NFK	N-formyl-kynurenine
NMDAr	N-methyl-D-aspartate receptor
OD_{600}	Optical density at 600 nm
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
PDB	Protein data bank
PEG x	Poly-ethylene glycol, average molecular weight of x
Pf	*Pseudomonas fluorescens*
PHBH	Para-hydroxybenzoate hydroxylase
PLP	Pyridoxal 5'-phosphate
PMSF	Phenylmethanesulfonyl fluoride

p-OHB	Para-hydroxybenzoate
QA	Quinolinic acid
ROS	Reactive oxygen species
rpm	Revolutions per minute
s	Seconds
Sc	*Saccharomyces cerevisiae*
SDS	Sodium dodecyl sulfate
$t_{1/2}$	Half-life
TAE	Tris-acetate-EDTA buffer
TDO	Tryptophan 2,3-dioxygenase
T_{eff}	Effector T cell
T_{reg}	Regulatory T cell
UV	Ultraviolet
Vis	Visible
w/v	Weight/Volume
w/w	Weight/weight
***Xc*TDO**	*Xanthomonas campestris* TDO

Contents

Chapter 1: Introduction

1.1. Overview of the kynurenine pathway

The overall function of the kynurenine pathway (KP) in mammals is the conversion of the essential amino acid L-tryptophan (L-Trp) into the catabolic oxidising agent nicotinamide adenine dinucleotide (NAD^+), as shown in Figure 1.1. This purpose was discovered in 1947 by Beadle *et al.* in a study on *Neurospora*[1]. The kynurenines themselves were discovered in 1904 by Ellinger who linked them to L-Trp in dog urine samples, hence their name[2]. In further studies it was found that 90 % of the L-Trp ingested enters the KP, while the other 10 % feeds into the serotonin and melatonin pathways and protein synthesis[3, 4]. Various equivalent proteins at the start of the KP are present in a number of bacteria including *Pseudomonas aeruginosa*, *Pseudomonas fluorescens*, and *Streptomyces antibioticus*[5]. These are associated with the production of quinolinic acid and anthranilic acid. Some bacterial species lack the aspartate to quinolinic acid pathway, including the order Flavobacteriales and genus *Xanthomonas*. These show genetic evidence for presence of the entire KP[5]. Several of the KP metabolites, such as L-kynurenine (L-Kyn), kynurenic acid (KA), quinolinic acid (QA) and 3-hydroxy-L-kynurenine (3-HK) are involved in a number of pathologies including cancer, Alzheimer's disease, acute pancreatitis, migraines and many more. The pathway involves numerous processes, enzymes and branching points; these factors increase the pathway's therapeutic potential and interest in the KP has grown accordingly.

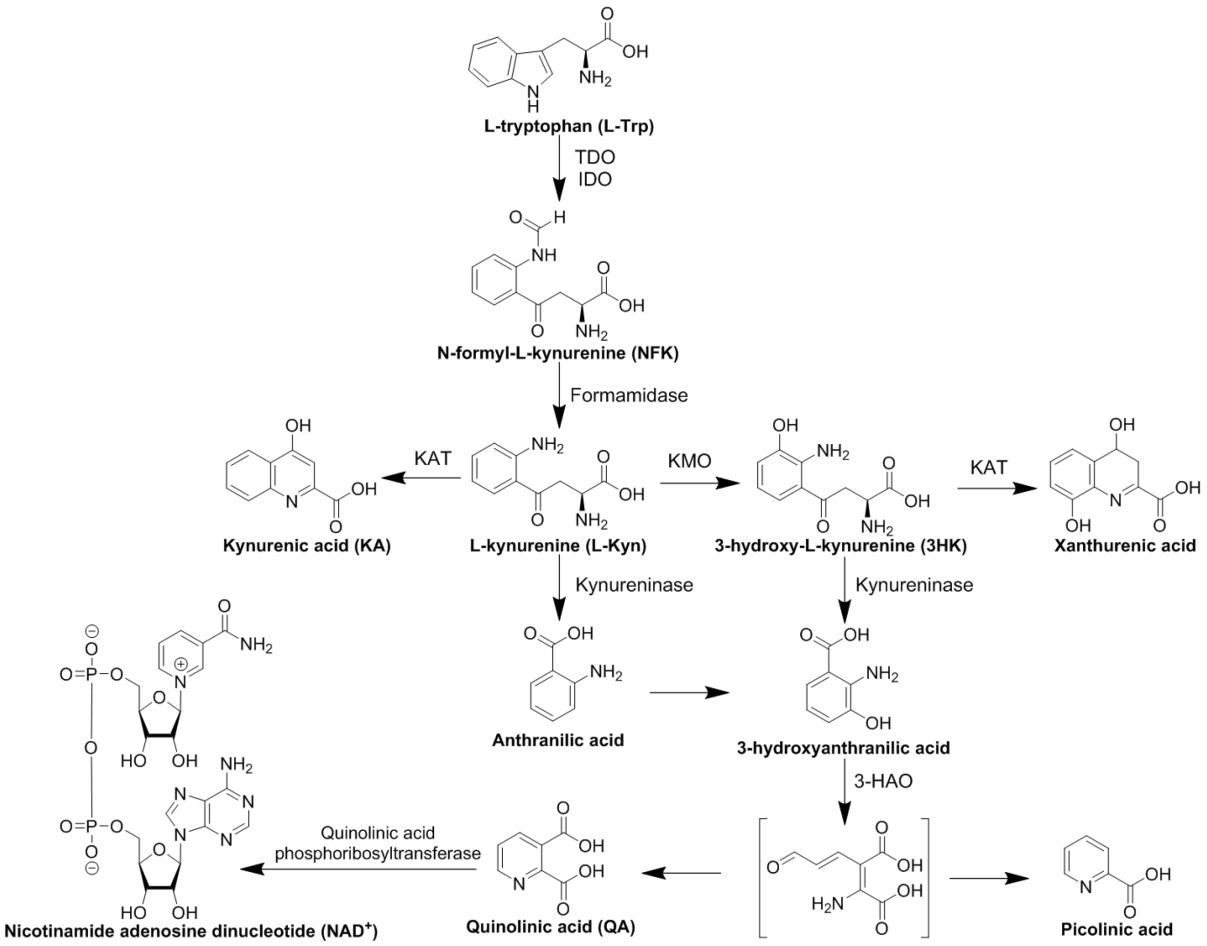

Figure 1.1 An overview of the kynurenine pathway (KP), steps with no named enzyme are non-enzymatic. Tryptophan 2,3-dioxygenase represented by TDO, indoleamine 2,3-dioxygenase represented by IDO, kynurenine aminotransferase represented by KAT.

The KP starts with the oxidation of L-Trp to N-formyl-kynurenine (NFK) by one of three enzymes; these are tryptophan 2,3-dioxygenase (TDO), indoleamine 2,3-dioxygenase 1 (IDO1) and indoleamine 2,3-dioxygenase 2 (IDO2). TDO is expressed predominantly in the liver, whereas IDO1 and IDO2 are expressed in various tissues[6-8]. NFK is rapidly hydrolysed to the more stable L-Kyn.

L-Kyn sits at one of the branching points of the KP; its downstream products are associated with both neurotoxic and neuroprotective roles depending on which path is followed. If converted by kynurenine aminotransferase (KAT) the resulting KA is a known neuroprotectant[9]. If converted by kynurenine 3-monooxygenase (KMO) the resulting 3HK is associated with neurodegeneration[10]. 3HK is further metabolised to form QA, which is also a known neurotoxin and this is eventually converted to the pathway's main product NAD^+. The later enzymes of the KP are found throughout the body in varying levels. The liver is the site of greatest expression for KMO and kynureninase but are found in many other tissues[11, 12]. KAT is expressed most in the brain and quinolinic acid phosphoribosyltransferase levels are highest in the kidney; both are expressed in various levels in other tissues[13, 14].

1.2. Metabolites of the kynurenine pathway

1.2.1. The L-tryptophan : L-kynurenine balance

The immune response is a powerful mechanism that, if left unchecked, can lead to debilitating and even fatal conditions[15]. Amongst the many systems that suppress the immune system, in order to protect self-cells, is the L-Trp : L-Kyn balance.

The initial conversion of L-Trp to NFK is the slowest, and therefore rate-limiting, step of the whole pathway. Upon formation NFK is rapidly converted to L-Kyn and as the KP's first stable metabolite a balance is formed between it and its precursor L-Trp. The ratio between the two plays a major role in immune response and is associated with a number of conditions relating to inflammation and immunology. This is modulated by L-Kyn production by the KP dioxygenases TDO, IDO1 and the recently discovered IDO2, and L-Kyn metabolism by KAT, KMO and kynureninase.

The expression of the KP dioxygenases in inflamed tissues was originally explained as a way of starving pathogens of L-Trp[16, 17]. However this theory did not take into account the effect of L-Trp starvation upon the immune system itself. The perceived role of the dioxygenases altered drastically from the immune response to immunosuppressant when it was discovered that IDO1 expression prevented allogeneic foetal rejection within mice[18]. The presence of IDO1 was seen to cause T-cell inhibition and has since been linked to regulation of many other immune related cells, such as dendritic cells (DCs), monocytes, macrophages, and microglia[19-21]. This led to further study on the role of the L-Trp : L-Kyn balance within immune suppression.

There are two theories as to how IDO1 produces such effects, one is L-Trp consumption and the other is the increase in kynurenine metabolites. Studies have suggested that both play a part in T-cell inhibition.

The lack of L-Trp has been shown to abort mitosis of T-cells, induce autophagy and stop T-cell proliferation[22, 23]. The increase in kynurenine metabolites such as 3-hydroxyanthranilic acid and QA suppresses the response from T-cells and even leads to T cell apoptosis [24, 25]. The alteration of the L-Kyn : L-Trp ratio also leads to naïve T-cells preferentially differentiating into regulatory T-cells, responsible for downregulating effector T-cells and reducing immune activity as a whole[26]. Evidence suggests this preferential differentiation is via aryl hydrocarbon receptor (AHR) activation[27, 28].

1.2.2. The kynurenic acid : quinolinic acid balance

The later steps of the KP involve another ratio between the two neuroactive metabolites KA and QA. The concentrations of both these metabolites are reliant upon availability of precursor molecules. This section focuses on the actions each metabolite has upon the nervous system and their implications regarding pathology.

1.2.2.1. Kynurenic acid

KA is produced by the conversion of L-Kyn catalysed by the pyridoxal 5'-phosphate (PLP) dependant kynurenine amino transferase isozymes (KATs 1-4)[29]. Four KAT isozymes have been reported in mammals; each require an α-ketoacid coenzyme to act as the amino acid acceptor[30]. In 1982 KA was shown to be neuroactive and has been implicated in both therapeutic and pathogenic roles in various diseases[31-33]. KA binds to a number of different receptors throughout the nervous system and other areas of the body, many of which are associated with disease. These include the N-methyl-D-aspartate receptor (NMDAr), the kainite- and α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid receptor (AMPAr), the α7 nicotinic acetylcholine receptor (nAChR) as well as the G protein-coupled receptor 35 (GPR35) and the aryl hydrocarbon receptor (AHR)[34].

The NMDAr is a subclass of glutamate receptor that regulates Ca^{2+} ion influx and it has long been associated with various neurological diseases. This can be through over-stimulation of the receptor causing neuron death, leading to neurodegenerative conditions such as Alzheimer's disease, Huntington's disease, and ischemic stroke. Subnormal activity of the receptor can also be detrimental and lead to schizophrenia and even neuronal apoptosis[35, 36]. Since KA is an antagonist it can act as a boon in some conditions and a bane in others. KA antagonism acts through competitive inhibition of the glycine co-agonist site of NMDAr with a K_i of 10 μM (IC_{50} = 36 μM)[37].

AMPAr is another subclass of glutamate receptor that requires a weaker stimulation than the NMDAr. When the AMPAr is stimulated this can trigger the stimulation of the NMDAr leading to Ca^{2+} influx[38]. At high (millimolar) concentrations KA acts as a competitive antagonist for AMPArs in a similar fashion to NMDArs. However, at low (submillimolar) concentrations KA has been shown to act as a "positive allosteric modulator", facilitating AMPAr responses to glutamate[39]. This presents KA as an AMPAr agonist at physiological levels and antagonist at higher levels[40].

Nicotinic acetylcholine receptors (nAChR) also regulate the influx of cations but are triggered by acetylcholine and have been linked to cognitive functions such as memory and learning[41]. Through modulation of inflammation, the stimulation of nAChR has been associated with a number of neurological diseases, such as Alzheimer's disease and schizophrenia[42]. KA is a non-competitive antagonist for nAChR with an IC_{50} much closer to physiological levels (submicromolar). This implies that the nAChR is the preferred target of endogenous KA[43].

KA has also been implicated in an immunological capacity. The AHR is a transcription factor involved in gene expression regulation leading to a number of responses associated with immunology and toxicology[44]. KA is a potent ligand for this receptor and elicits a response at physiological levels (IC_{50} of 1.4 μM), and is therefore believed to be one of the few endogenous ligands of AHR[45]. The response to KA binding to AHR is the expression of interleukin-6, a mediator of inflammation and immune

response[45]. The AHR response to L-Kyn is also relevant in immune escape of cancer, which is further discussed in section 1.3.1.

When at high concentrations KA also stimulates GPR35, a G protein-coupled receptor predominantly expressed in immune cells. This and the heightened levels of KA during inflammation gave rise to the theory that the two could play an important role in immunological modulation[46, 47]. However the discovery of GPR35 expression within the central and peripheral nervous systems has reduced the significance of KA's possible role in this area[48, 49].

1.2.2.2. Quinolinic acid

Quinolinic acid (QA) is the last precursor of the KP before NAD^+ formation, and is mainly produced in the microglia[50]. The route to QA production within the KP starts at the branch point of L-Kyn, as shown in Figure 1.1[51]. KMO converts L-Kyn to 3HK, which is cleaved by kynureninase to produce 3-hydroxyanthranilic acid (3-HA). This is then converted to aminocarboxymuconate semialdehyde by 3-hydroxyanthranilic acid oxygenase (3-HAO). This unstable intermediate cyclises non-enzymatically through condensation of the aldehyde with the amine to form QA[52].

QA was discovered to be neuroactive when it caused convulsions in mice after being injected into the brain[53]. It has since been established as a potent excitotoxin of the central nervous system (CNS) predominantly as an agonist for the NMDAr[54]. The uptake of QA within the synapse is relatively slow; this causes the activation of NMDAr to occur for an extended period of time, which in turn allows large levels of Ca^{2+} ions to enter the cell. This, as discussed above, leads to neuron death and is implicated in many neurodegenerative diseases[55]. QA neurotoxicity is most prevalently observed in the hippocampus and striatum due to the high levels of NMDArs situated there[56, 57].

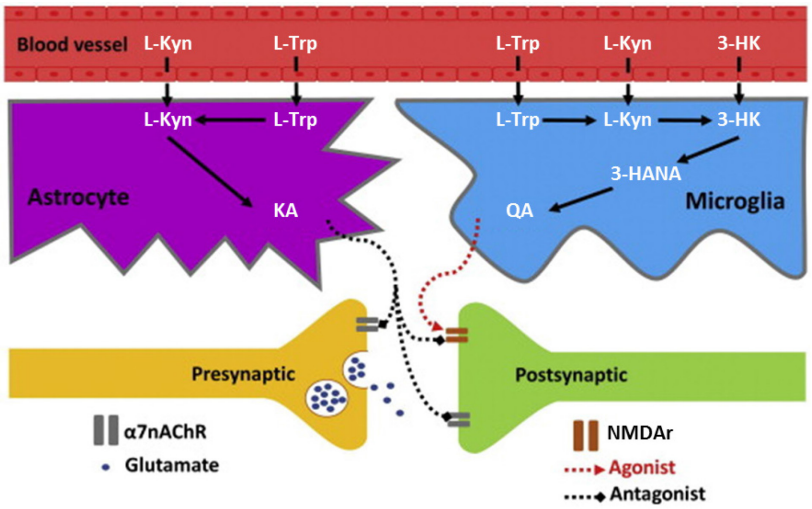

Figure 1.2 Schematic of KA and QA within the nervous system and their effect on NMDAr and α7nAChR. Figure adapted from Maddison *et al*.[58]

QA toxicity is also due to oxidative stress through various mechanisms in conjunction with, or independent of, the NMDAr. QA has been shown to increase nitric oxide synthase activity, producing free radicals and, through complexation with Fe^{2+}, create reactive oxygen species (ROS), thus inducing lipid peroxidation[59, 60]. QA even acts as a double-edged sword, decreasing the ability of anti-oxidants such as glutathione and superoxide dismutase to deal with oxidative species[61].

1.2.3. 3-hydroxy-L-kynurenine

3-hydroxy-L-kynurenine (3HK) is the direct oxidation product of KMO and is mostly found in the lens of the eye where it acts as a UV filter[62]. Studies suggest the levels of 3HK in the brain are low under normal conditions but are elevated under pathological conditions[10, 63]. 3HK has been assigned in the literature as both a beneficial antioxidant and a deleterious oxidative toxin. 3HK is capable of donating electrons and scavenging radicals, suggesting a role as an endogenous neuroprotectant[64, 65]. In more oxidative conditions, however, 3HK is capable of promoting ROS production in much the same manner as QA, leading to neuron death and oxidative stress[66]. Most of the evidence suggests that damage is caused when levels of 3HK are above normal physiological levels, advocating the case for 3HK as an endogenous antioxidant[62]. Under pathological conditions, however, the rise in 3HK concentrations may reach a tipping point towards harmful action.

1.3. The kynurenine pathway in disease

1.3.1. Dioxygenases and cancer

1.3.1.1. The L-Tryptophan : L-Kynurenine balance

As mentioned in section 1.2.1 the L-Trp : L-Kyn ratio is inherently important in immune response modulation affecting T-cell growth, activity and differentiation. This is achieved through both decreasing L-Trp and increasing L-Kyn levels. While useful in preventing autoimmune reactions this system can also be beneficial to cancer cells. The expression of KP dioxygenases in order to perturb the L-Trp : L-Kyn balance can allow cancers to evade the immune system.

Immune evasion is now seen as one of the "hallmarks" of cancer and different tumours exhibit a myriad of mechanisms in order to achieve it. Three common characteristics of immune modulation by cancers are evasion of immune recognition, chronic inflammation in the microenvironment and suppression of immune action[67]. The KP dioxygenases are predominantly involved in the latter.

1.3.1.2. Expression of dioxygenases

Constitutive expression of IDO1 in tumour cells was discovered in many human cancers in 2003; including prostatic, colorectal, pancreatic and cervical carcinomas[68]. The expression of TDO has also been identified in cancers such as bladder, colorectal, and hepatic carcinomas as well as glioblastomas[69]. The expression of IDO2 in some cancers has been reported. However, these have been disputed citing the presence as inactive IDO2 polymorphs; this is discussed further in section 1.4.3. The contradictions found within the literature add to the uncertainty of IDO2's role[70].

1.3.1.3. Action of the aryl hydrocarbon receptor

As the concentration of L-Kyn increases around cancer cells this leads to the activation of the AHR[28]. The AHR is a transcription factor that is regulated by planar aromatic hydrocarbons and L-Kyn activates AHR with an EC_{50} in the micromolar range (12-37 μM)[71]. When activated it translocates to the nucleus and binds to specific DNA strands thus activating certain genes. Under normal conditions activation of AHR has a wide range of results including cell development, cell migration and immune response. This can have a number of consequences relating to immunosuppression and immune escape of tumours.

The positioning of the AHR dictates its action, as seen in Figure 1.3. The AHR can be on the cancer cell itself, on effector T cells (T_{eff}) and on dendritic cells (DC)[71].

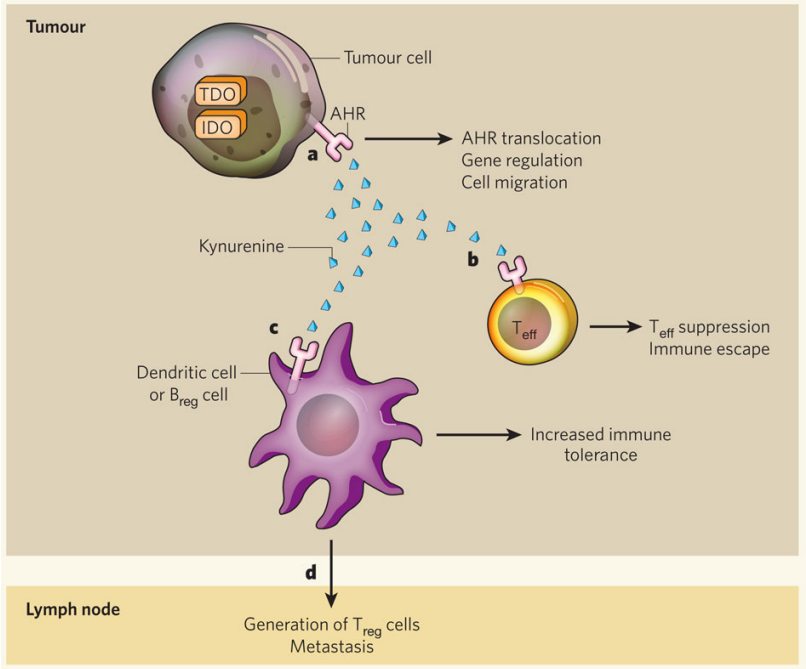

Figure 1.3 Schematic showing effect of L-Kyn activated AHR depending on the position of the AHR. Image adapted from Prendergast[72].

When AHR is activated on T_{eff} cells their activity against cancer cells is reduced. In addition to this; levels of proliferation, activation and growth are also decreased[71, 73].

When AHRs located on naïve T-cells are activated by L-Kyn this leads to the generation of regulatory T-cells. These reduce the activity of T_{eff} cells, lowering the immune response towards the cancer. This effect is increased when the AHR on DCs is activated. DCs are well known to play a role in naïve T-cell differentiation, and AHR activation causes the further expression of IDO1 and IDO2 by DCs. This in turn increases L-Kyn levels generating further T_{reg} cells[28, 74]. If these DCs travel from the microenvironment surrounding the cancer to lymph nodes this can increase systemic tolerance, making metastasis more likely.

When situated on the cancer itself AHR activation induces expression of genes that promote both cell growth and motility. This increases the likelihood of metastasis and malignancy[71].

The expression of KP dioxygenases, especially IDO1, has been shown to worsen a cancer patient's prognosis significantly[75]. The knockdown of IDO1 in mice using both genetic and chemical based inhibition was shown to reduce cancer growth and increase effectiveness of anti-cancer drugs[68, 76]. IDO1 is therefore an attractive drug target and several compounds are being investigated as potential IDO1 inhibitors, some of which are in clinical trials, including the hydroxyamidine Epacadostat and the imidazole NLG919[77].

A clinical trial of 1-methyl-D-tryptophan (1-DMT), shown in Figure 1.4, was started in 2007[78]. The mechanism of action surrounding 1-DMT is disputed. 1-DMT exhibits no inhibition of IDO1 *in vitro* unlike its enantiomer 1-methyl-L-tryptophan (1-LMT)[79]. The literature is divided in the suitability of 1-DMT as a potential drug. Its effects have been posited as the result of side reactions within the cell rather than direct action upon IDO1[80]. Further uncertainty stems from the two different assays that have been used throughout the years, which are further discussed in section 1.4.3.

More recent work by Tomek *et al.* evaluated the National Cancer Institute Diversity III library as IDO inhibitors[81]. The most promising compound, "pyrimidinone 4", depicted in Figure 1.4, showed effective IDO1 inhibition (IC_{50} value 4.4 ± 1.5 μM) and low cytotoxicity (LC_{50} > 117 μM).

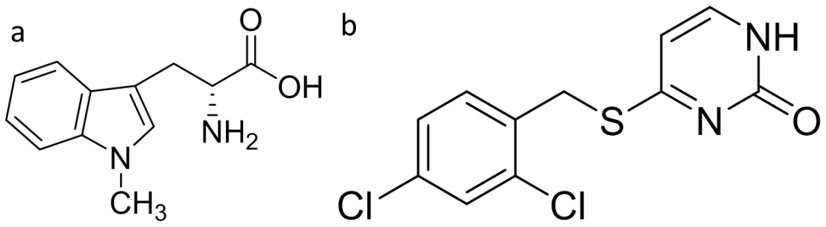

Figure 1.4 a. Structure of 1-methyl-D-tryptophan (1-DMT). **b.** Structure of "pyrimidinone 4".

1.3.2. Kynurenine 3-monooxygenase and disease

As mentioned in section 1.2.2 an alteration in precursor concentrations leads to a change in the balance between KA and QA. An increase in KA levels sees a drop in QA levels and vice versa. L-Kyn is seen as the branch point and has the potential to be converted to both of these metabolites. KMO catalyses the first step on the way to QA production. Inhibition of KMO causes a decrease in concentrations of both 3HK and QA; this in turn increases KA levels[82, 83]. Therefore inhibition of KMO is seen as a viable therapeutic strategy[34, 84, 85].

This section describes the role KMO plays in a number of diseases as well as the effect KMO inhibition has had upon models of these diseases.

1.3.2.1. Huntington's disease

The neuroactivity of KP metabolites has led to the association of KMO with a number of neurological diseases[84]. Some of the most extensive work surrounding KMO is focused on its involvement in Huntington's disease (HD). HD is a fatal genetic disease that causes neuron death due to the misfolding of the huntingtin protein[86]. The connection between HD and the KP was made through the similarities found in symptoms of rats with HD and those injected with QA[87]. Analysis of QA and KA concentrations soon corroborated this theory. KA levels are reduced in the cerebral cortex of HD mouse models and HD patients, while separate studies showed levels of QA to be elevated in HD mouse models and early-stage HD patients[88-91]. The impairment of mitochondrial function that is associated

with HD hinted towards KMO's involvement and the first evidence of KMO's clinical potential came in 2005[92]. The deletion of the gene coding for KMO from yeast suppressed the toxic effects of a mutant huntingtin fragment in yeast while also eliminating QA[93]. Since then the knockdown of KMO has also shown reduction in QA levels and symptoms of HD in both mice and fruit fly models[85, 94].

1.3.2.2. Alzheimer's disease

Another neurological disease associated with the KP is Alzheimer's disease (AD)[95]. AD causes plaques of misfolded peptides to form in the brain, as well as the phosphorylation of tau proteins[96, 97]. High 3HK and QA levels are present in AD patients and QA increases symptoms of AD such as tau phosphorylation and the expression of AD associated genes[55]. Chemical inhibition of KMO in mice saw an increase in concentrations of KA, which slowed the progression of the disease and ameliorated the symptoms presented[84]. In a more recent study the inhibition of both KMO and TDO led to reduction of symptoms of AD in fruit flies[98].

1.3.2.3. Stroke

Due to its involvement in oxidative stress and neuronal damage the KP was investigated in its role in ischemic stroke. The levels of KP enzymes were found to be elevated in the hippocampus of post-ischemic mice, as was the level of QA in the brain overall[99]. Chemical inhibition of KMO reduced the amount of QA and 3HK in post-ischemic gerbil brains, and increased levels of KA. This reduced neuron death and decreased glutamate levels thus reducing oxidative damage[100].

1.3.2.4. Acute pancreatitis

Recently KMO has been implicated in acute pancreatitis (AP) and the multiple organ failure that often follows[101]. The presence of 3HK is increased in mesenteric lymph tissue in rats, which increases cytotoxicity towards the lungs. The level of L-Kyn is also elevated in the blood of both rats and humans with AP[101]. In a collaborative study with GlaxoSmithKline, the use of KMO inhibitor GSK180 on AP model rats significantly altered the levels of KP metabolites in the blood stream. This provided effective protection against multiple organ failure[11].

The KP is associated with numerous other neurological conditions including Parkinson's disease, amyotrophic lateral sclerosis, schizophrenia and even migraine[34, 102-104]. Many studies of these conditions have shown the modulation of the KP to produce a positive effect. KMO therefore presents itself as an attractive drug target for numerous conditions.

1.4. Dioxygenases of the kynurenine pathway

1.4.1. Tryptophan 2,3-dioxygenase

TDO is a cytosolic homotetrameric b-type heme protein found most readily in the liver. It was first purified in 1936 and has high specificity towards L-Trp[105]. Its primary role is in regulation of the L-Trp :

L-Kyn ratio on a systemic scale. Crystal structures of *Xanthomonas campestris* TDO (*Xc*TDO) and *Ralstonia metallidurans* TDO have revealed the enzyme to be more of a "dimer of dimers" given the N-terminal residues form part of the active site with the adjacent subunit, as shown in Figure 1.5[106, 107]. Since then *Drosophila melanogaster* TDO and human TDO structures have also been solved in 2013 and 2014 respectively[108, 109]. The human enzyme structure, however, is of the apo protein and is without heme.

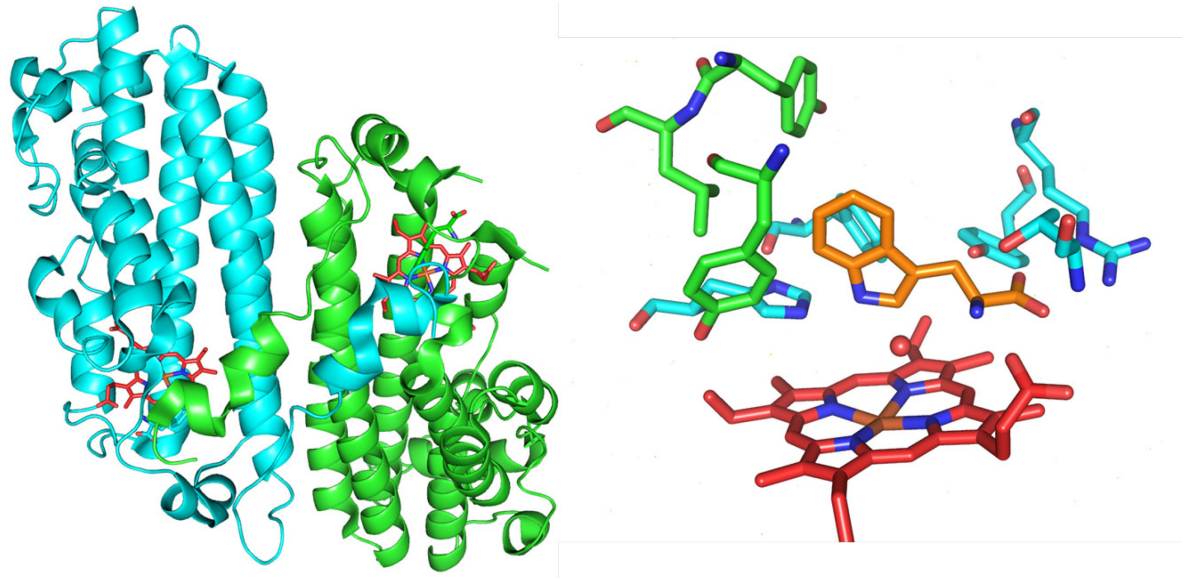

Figure 1.5 Crystal structure of *Xc*TDO in complex with heme and tryptophan. Left shows half of the tetramer, i.e., one dimer of dimers. Right shows L-Trp (orange) interacting with active site residues and heme (red). Residues from the N-terminus of the adjacent monomer are shown in green (PDB: 2NW8)[106].

1.4.2. Indoleamine 2,3-dioxygenase 1

IDO1 is a monomeric b-type heme protein found throughout the body except the liver. Isolated in 1967 IDO1 is more promiscuous than TDO, oxygenating various substrates in addition to L-Trp, including D-Trp, 1-Me-L-Trp, 5-F-Trp, serotonin and trytpamine[110]. IDO1 is expressed at lower levels than TDO, is present throughout the body, and its expression is strongly activated by interferon gamma (IFN-γ), a signalling molecule associated with immune response and inflammation. Also the K_M for L-Trp of IDO1 (0.021 μM) is significantly lower than TDO (0.19 μM)[111]. It is believed therefore that IDO1 regulation of the L-Trp : L-Kyn ratio is achieved within the local microenvironment in response to pathogenesis or injury[112]. Several crystal structures of the human IDO1 have been solved with inhibitors bound, showing two clear domains within the protein, as seen in Figure 1.6. The larger, all helical, domain is made of 13 α-helices and two 3_{10} helices. This domain makes up most of the heme-binding pocket. The smaller C-terminal domain is formed of six α-helices, two β-sheets, and three 3_{10} helices. Much like the N-terminal residues of TDO the small domain makes up one side of the heme-binding pocket[113].

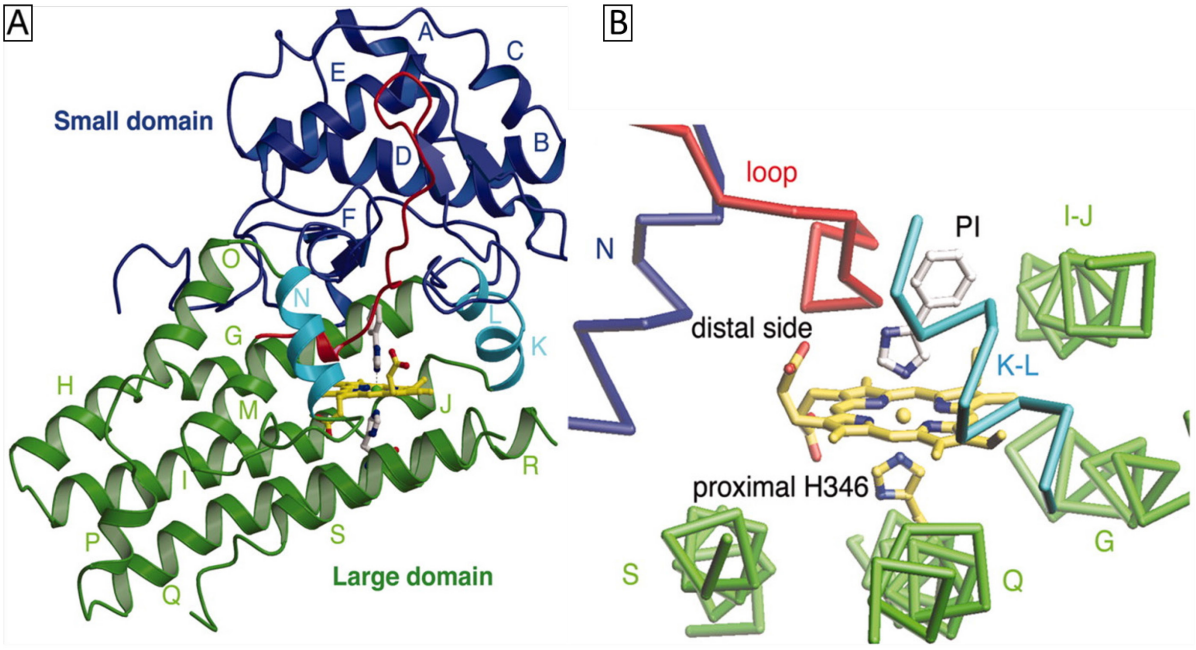

Figure 1.6 Crystal structure of hIDO1 in complex with heme and 4-phenylimidazole (PI). **A.** Shows overall protein structure. **B.** Shows heme binding site with contributions from both small and large domains. Adapted from Sugimoto *et al*. (PDB: 2D0T)[113].

1.4.3. Indoleamine 2,3-dioxygenase 2

IDO2 was discovered in 2007 and is much less well characterised than the other dioxygenases in the KP[114]. The gene for IDO2 is located very close to the IDO1 chromosome, making it a likely genetic predecessor of IDO1[114]. IDO1 and IDO2 share a BLAST sequence homology of 44%; an alignment of the two sequences is shown in Figure 1.7[115].

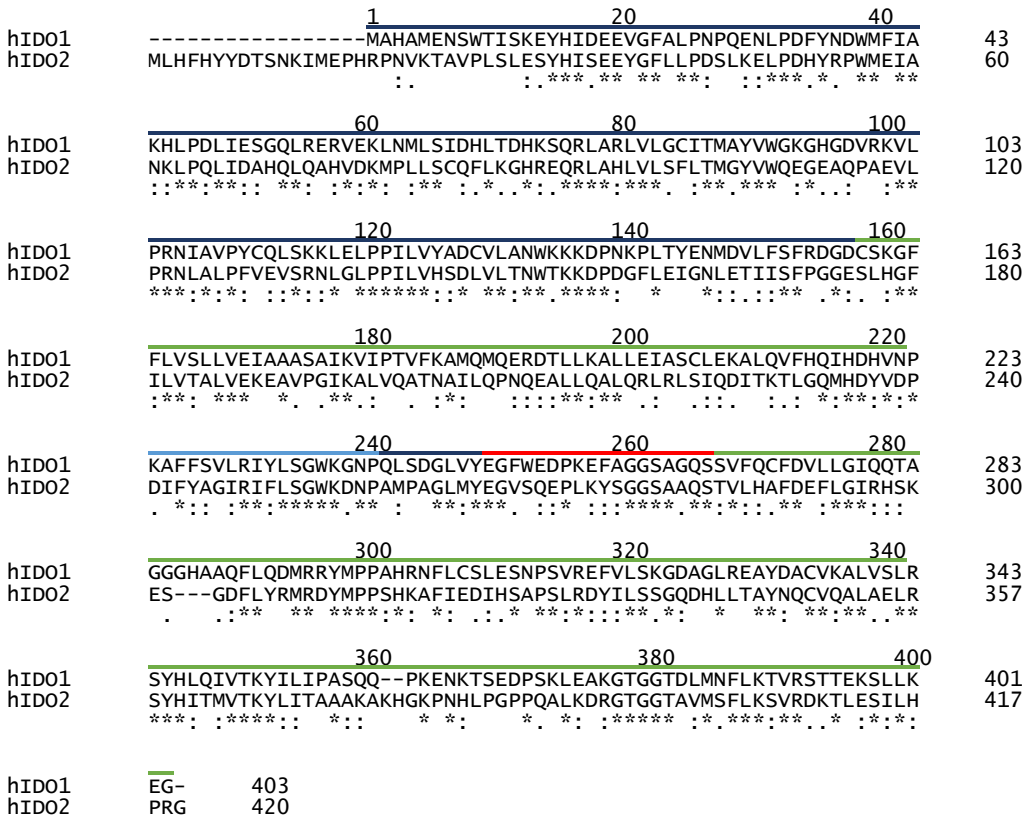

Figure 1.7 Sequence alignment comparing hIDO1 and hIDO2. Alignment generated using Clustal Omega[116]. The Uniprot identifier codes are hIDO1 - P14902, hIDO2 - Q6ZQW0. hIDO1 domains are represented through colours; the small domain is dark blue, the large domain is green, residues connecting the domains are red and light blue.

IDO2 is expressed in some of the tissues that express IDO1, but unlike IDO1 is also found in the liver[117]. Activation of IDO2 expression by IFN-γ is less clear, with partial up-regulation occurring in human cells but not in other models such as mice[118]. IDO2 is inherently more hydrophobic and therefore less stable than IDO1 and this has led to difficulties isolating and analysing it. The enzyme, while able to catalyse the dioxygenation of L-Trp, is much less effective than IDO1 causing more difficulties in kinetic analysis. Activity of IDO1 and IDO2 has been examined using different assays, one based on methylene blue (MB), the other based on cytochrome b_5 (cyt b_5)[119]. A comparison of the kinetic properties of the two enzymes is shown in Table 1.1.

	mIDO1	mIDO2
K_M MB assay	28 ± 4 μM	12000 ± 3000 μM
k_{cat} MB assay	2.6 ± 0.1 s^{-1}	0.02 ± 0.002 s^{-1}
K_M cyt b_5 assay	29 ± 9 μM	530 ± 100 μM
k_{cat} cyt b_5 assay	0.09 ± 0.008 s^{-1}	0.10 ± 0.008 s^{-1}

Table 1.1 A comparison of mouse IDO1 (mIDO1) and mouse IDO2 (mIDO2) Michaelis-Menten kinetic properties for the substrate L-Trp through two different assay systems[119].

The disparity in results for IDO2 between the MB and cyt b_5 assays, suggests either an effective assay has yet to be found for IDO2 or that L-Trp is not the natural substrate for IDO2. This in turn would raise questions regarding the physiological function of IDO2. Another complication of IDO2 is the presence of its genetic polymorphs. These polymorphs either drastically reduce or totally cease IDO2 activity and may be found in up to 50% of the population[117]. The high-resolution structure of IDO2 has yet to be solved leaving many questions about this refractory molecule unanswered.

1.4.4. The dioxygenase mechanism

The proposed chemical mechanism of the KP dioxygenases is believed to be conserved, but literature descriptions have been developed over the years as more evidence has come to light[120]. The initial step was proposed as a base-catalysed proton abstraction, as seen at position 1 in Figure 1.8[121]. This was supported by the fact that the inhibitor 1-methyl-tryptophan (1-MT), shown in Figure 1.4, was unable to undergo base abstraction since it lacked the necessary proton. The abstraction was then followed by the formation of either a Criegee or dioxetane intermediate.

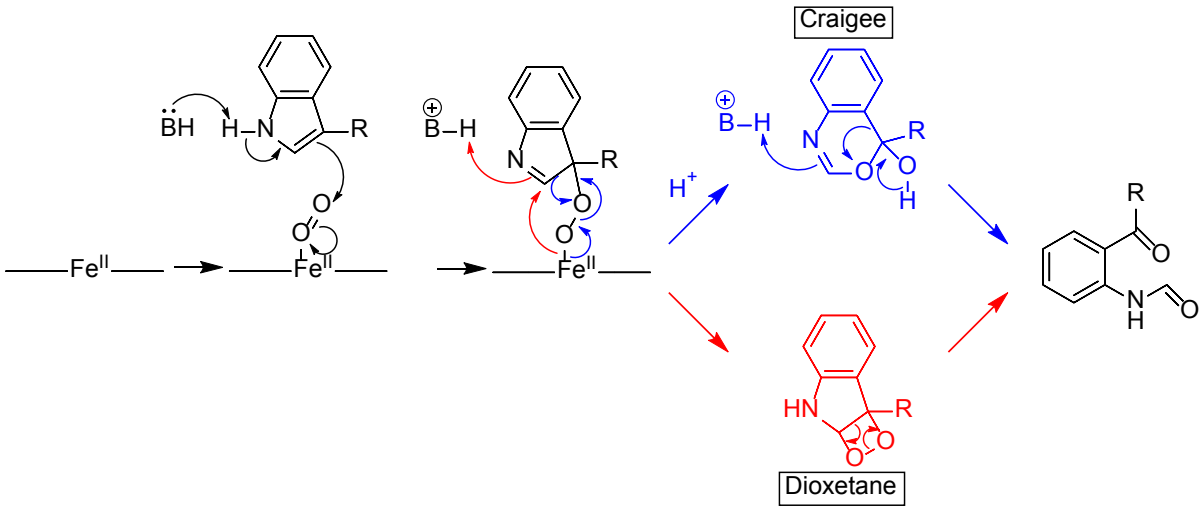

Figure 1.8 Initially proposed mechanism of KP dioxygenases. The heme plate is shown as a line surrounding Fe. The Criegee route is shown in blue, the dioxetane route is shown in red[120].

However, when it was established that 1-MT is in fact a slow substrate, and the crystal structure of IDO1 showed the active site lacking any suitable base residue candidate, the base abstraction theory was abandoned[122]. The formation of a transient Compound II ferryl species, shown in Figure 1.9, was suggested through computational modelling and Raman data[123-125]. The presence of this ferryl species was confirmed by mass spectrometry when IDO1 turnover of several L-Trp analogues was studied. The presence of the appropriate ferryl species in all the L-Trp analogues indicates a common mechanism[126]. The formation of the ferryl species has been proposed via electrophilic addition and via radical addition; the steps following the formation of the ferryl species are based on QM/MM calculations; see Figure 1.9 [127, 128].

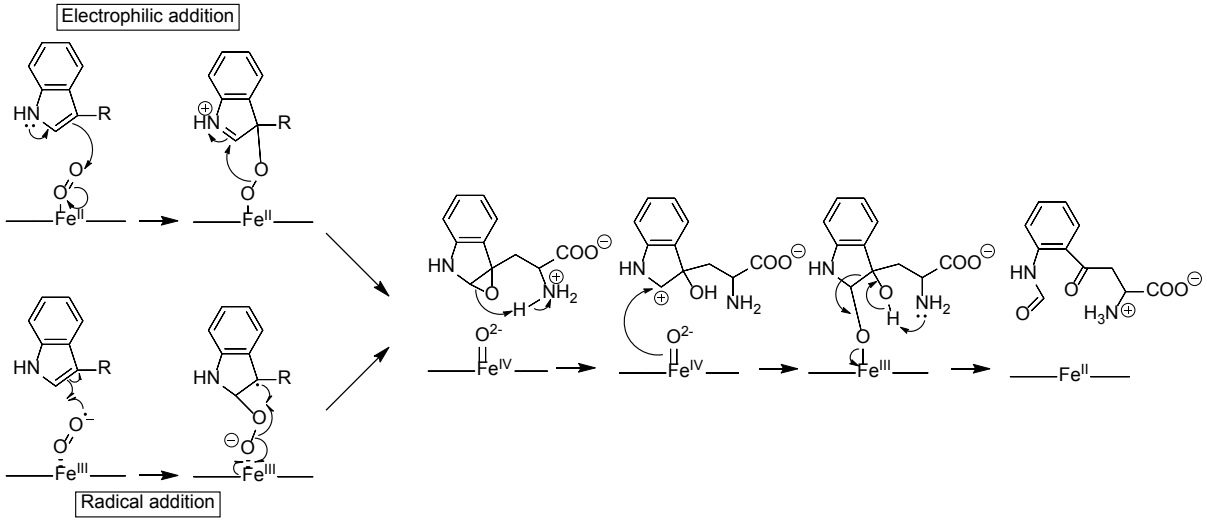

Figure 1.9 Current mechanism for TDO and IDO1. The heme plate is shown as a line surrounding Fe. The electrophilic addition route is shown on top, the radical addition route is shown on the bottom[127, 128].

1.5. Kynurenine 3-monooxygenase

KMO is a type A flavoprotein aromatic hydroxylase that catalyses the conversion of L-Kyn to 3HK. The human form of KMO has a mass of approximately 55 kDa and is located in the outer membrane of the mitochondrion. It is predominantly expressed in the liver and kidney but is also expressed in microglia and infiltrating macrophages within the brain[129, 130].

The C-terminus of KMO binds to the outer mitochondrial membrane and is therefore very hydrophobic. This greatly reduces the stability of recombinantly expressed human KMO (hKMO) and therefore obtaining hKMO has been challenging, with early attempts only producing partial purifications of inactive enzyme[131, 132]. The C-terminus, due to its hydrophobicity, aggregates with other proteins and causes co-elution during purification attempts. Due to these problems, KMOs from other organisms have been used to study its mechanism and structure. KMO lends itself to these studies due to the high

level of conservation between KMOs of different species, especially in terms of the active site, as can be seen in Figure 1.10.

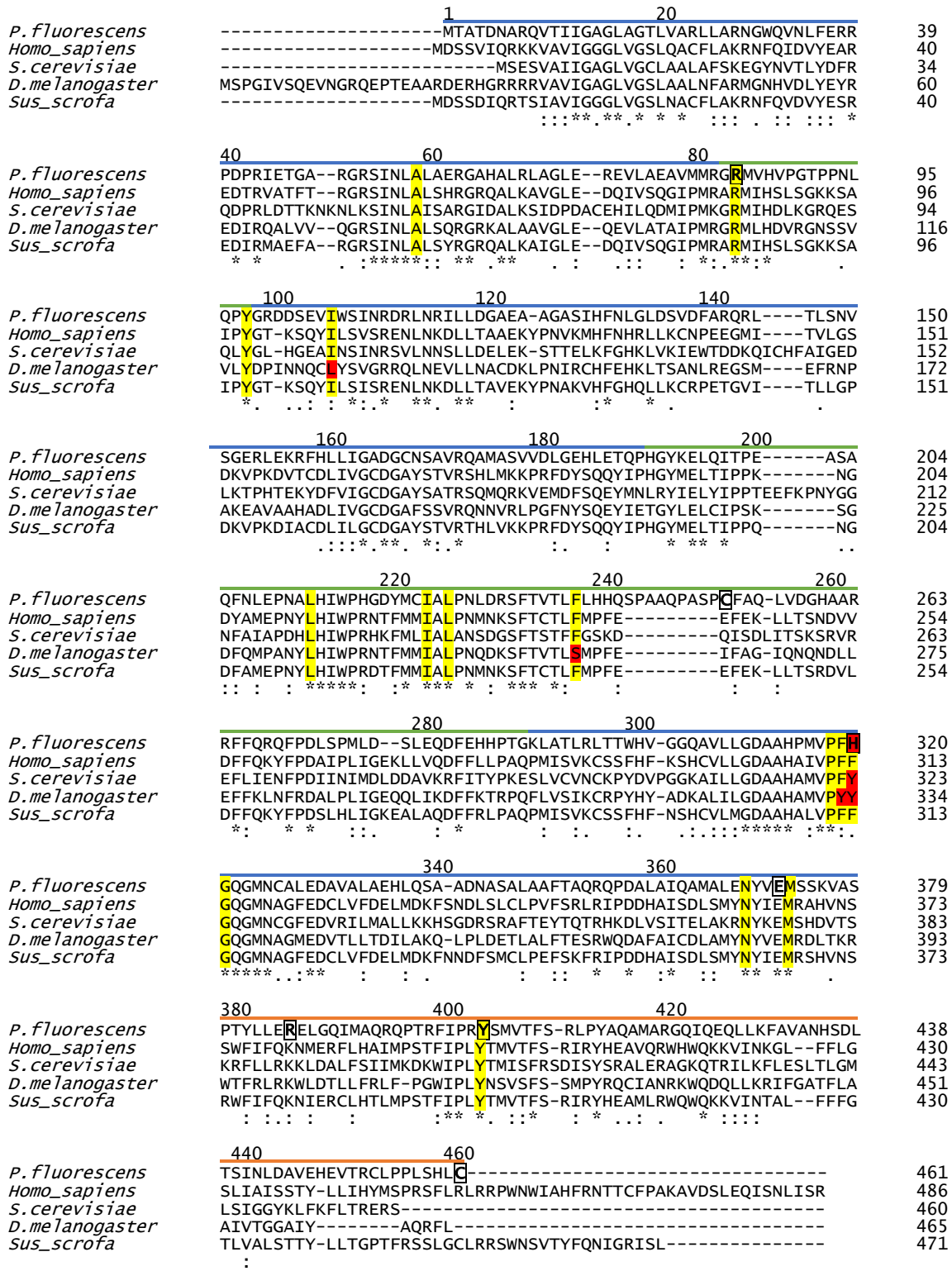

Figure 1.10 Sequence alignment comparing KMO from various species. The active site residues are highlighted (yellow for conserved, red for non-conserved). Encircled *Pf*KMO residues in bold are sites of mutation that are presented throughout this thesis. These include R84, Y404 and H320 in chapter 4 and R386 and E372 in chapter

5. C252 and C461 are discussed in chapter 1. *Pf*KMO domains are represented through colours; blue is the FAD-binding domain, green is a domain conserved in flavoprotein monooxygenases, orange is a domain unique to KMO[133]. Alignment generated using Clustal Omega[116]. The Uniprot identifier codes are *Saccharomyces cerevisiae* – P38169, *Pseudomonas fluorescens* - Q84HF5, *Drosophila melanogaster* - A1Z746, *Homo sapiens* - O15229, *Sus scrofa* - Q9MZS9.

1.5.1. Mammalian kynurenine 3-monooxygenase

One of the first models for hKMO was pig KMO, which also suffered from co-elution with other unwanted proteins due to its hydrophobic C-terminus[134]. In attempting to solve this issue Hirai *et al.*[135] altered the KMO gene to truncate the C-terminus. The solubility of KMO increased but the activity was totally abolished. By altering truncation positions, that both increased and decreased activity, the last 20 amino acids of KMO were identified as responsible for membrane localisation of the protein. The C-terminus was also recognised as playing a significant role in the activity of the enzyme.

Later work has expressed human KMO recombinantly using human embryonic kidney (HEK) and Chinese hamster ovary (CHO-1) cells with no purification attempted[136, 137]. The expression of hKMO in *E. coli* was reported in 2014 and with the addition of a hydrophilic FLAG tag even resulted in active hKMO in the soluble fraction[138]. The resulting purification process, however, still resulted in co-elution with other proteins. When the C-terminus was truncated the activity of the protein was lost, matching the result of pig KMO. In another study a glutathione S-transferase-KMO fusion protein was expressed in COS-1 cells, and was successfully purified to a purification factor of 100[139].

1.5.2. Non-mammalian kynurenine 3-monooxygenase

Research on non-mammalian KMO has been more successful. KMO from *Saccharomyces cerevisiae* (*Sc*KMO) was successfully expressed recombinantly in *E. coli* and purified[140]. This was also a truncated form of the enzyme lacking the C-terminus. Unlike the pig or human KMO however, the activity of the enzyme remained unaffected. This led to the first structure of KMO being published, which is further discussed in section 1.6.

Much work has been conducted on KMO from *Pseudomonas fluorescens* (*Pf*KMO)[11, 141-143]. The main advantage of *Pf*KMO is that being prokaryotic it is not membrane bound. The C-terminus, therefore, is much less hydrophobic and the protein is more stable overall. *Pf*KMO has a 36% sequence identity to hKMO and the construct (kindly gifted by Graham Moran, University of Minnesota) is used throughout the work presented here. *Pf*KMO was used to propose the mechanism of KMO and was recently used in a study on KMO inhibition as a treatment for multiple organ failure due to acute pancreatitis[11, 142].

1.5.3. Para-hydroxybenzoate hydroxylase

As previously mentioned KMO is a type A flavoprotein aromatic hydroxylase. This means the protein is encoded by one gene, contains one dinucleotide binding domain (Rossman fold), tightly and non-covalently binds an FAD cofactor, depends upon NADPH as a coenzyme, and $NADP^+$ is immediately released upon FAD reduction[144]. The structure of FAD is shown in Figure 1.11. A number of well-studied type A flavoprotein aromatic hydroxylases have been used to direct the elucidation of KMO's mechanism; the best-characterised of these is para-hydroxybenzoate hydroxylase (PHBH).

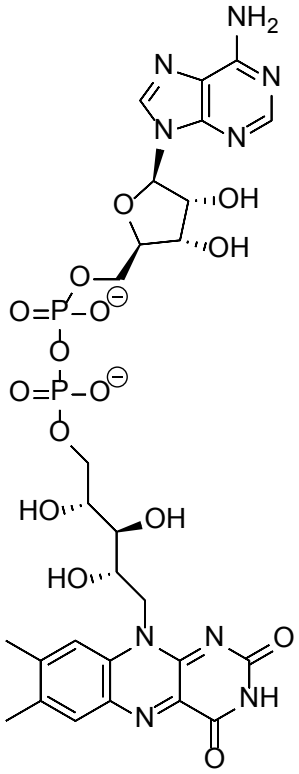

Figure 1.11 Structure of flavin adenine dinucleotide (FAD).

PHBH is an enzyme that degrades aromatic compounds by adding an OH group, thus allowing further catabolism. It is an ideal model enzyme for the family of type A FAHs as it can be produced in high levels, is easily purified and is very stable[145]. The formation of FAD intermediate species throughout the reaction can also be easily followed using UV and fluorescence spectroscopy. This work produced a mechanism that became the model for flavoprotein monooxygenases.

1.5.3.1. Mechanism of para-hydroxybenzoate hydroxylase

The proposed mechanism is formed of two parts, a reductive half involving NADPH and an oxidative half involving O_2[145]. The enzyme binds both NADPH and the substrate p-hydroxybenzoate (p-OHB) rapidly. This is followed by the rapid reduction of FAD, but this only occurs in the presence of the substrate p-OHB[146]. The release of $NADP^+$ is relatively slow and is comparable to the overall reaction

rate. At this point in the reaction the undesirable release of p-OHB is very slow; this is to ensure that the reduced FAD does not react with O_2 and produce toxic H_2O_2. The oxidative half of the reaction is very rapid, with O_2 reacting to form flavin-C4a-hydroperoxide. This reacts with the substrate via electrophilic substitution to form the product 3,4-dihydroxybenzoate and flavin-C4a-hydroxide; this decays to FAD and water. The last step is release of product and water; water release is much slower than the product and is believed to be responsible for determining the overall reaction rate. A schematic of the main steps of the mechanism is seen in Figure 1.9.

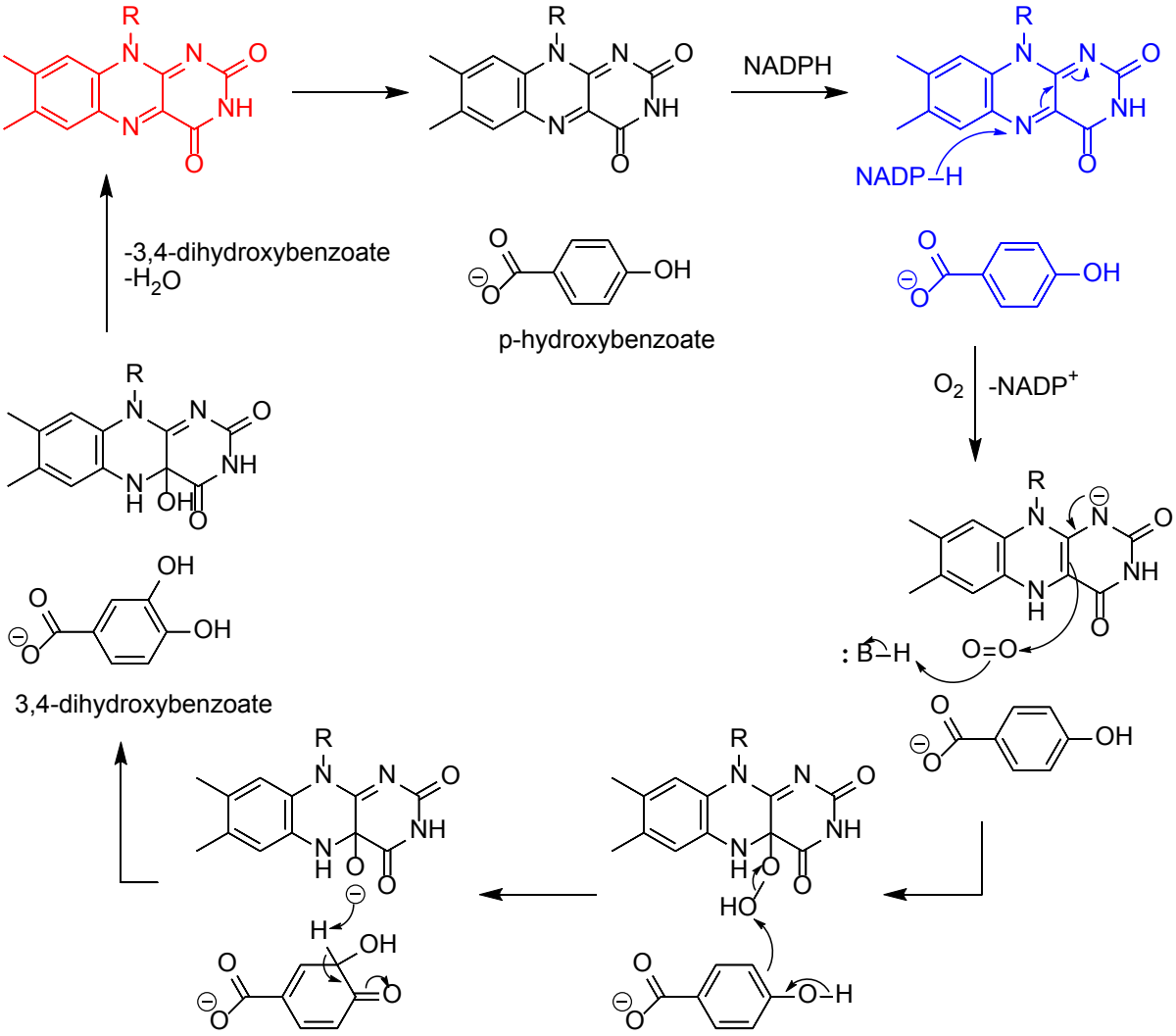

Figure 1.12 Schematic of the mechanism of PHBH. Stages show which conformation the protein is in, red is *open*, black is *in* and blue is *out*. Here and for the remainder of the thesis, the R group on the FAD isoalloxazine ring represents the rest of the dinucleotide. Mechanism proposed by Entsch *et al.*[145], adapted from Cole *et al.*[147].

1.5.3.2. Structural conformations of para-hydroxybenzoate hydroxylase

Throughout the stages of the reaction the protein cycles through a number of conformational states. First is the *open* conformation; this forms a path from bulk solvent to the interior of the protein, allowing for binding and release of substrate and product[147, 148]. The protein shifts to the *closed* position in which

catalysis occurs. This conformation closes the path to the interior and excludes H_2O from the active site ensuring no formation of H_2O_2 occurs. Further conformational changes within the active site are necessary for the reductive and oxidative half reactions. During the reductive half-reaction NADPH in the solvent requires access to the isoalloxazine of the FAD in order to reduce it, the protein is however in the *closed* conformation. This is overcome by the isoalloxazine moving 7-8 Å around the C-1' to C-2' bond towards the surface of the enzyme, known as the *out* position. The movement is controlled by the presence of p-OHB increasing specificity towards p-OHB as a substrate. This movement is shown in Figure 1.14, depicting PHBH in complex with the product analogue 2,4-dihydroxybenzoate. The structure of 2,4-dihydroxybenzoate is shown in Figure 1.13.

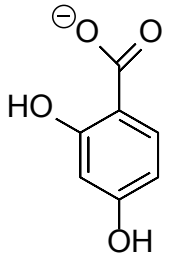

Figure 1.13 Structure of 2,4-dihydroxybenzoate

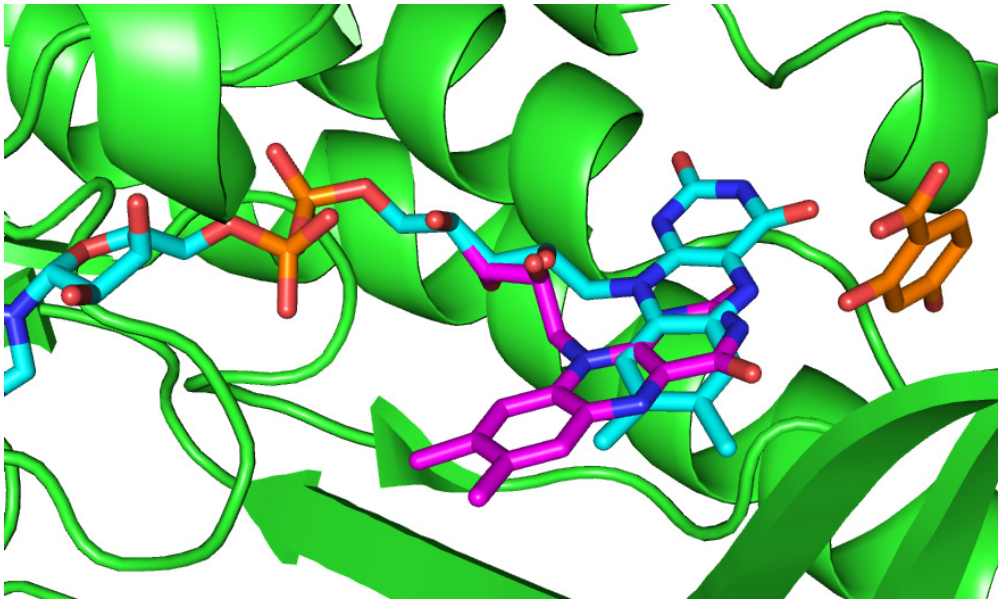

Figure 1.14 Crystal structure of PHBH from *Pseudomonas aeruginosa* complexed with 2,4-dihydroxybenzoate (orange) at the active site. FAD is shown in both *in* (cyan) and *out* (magenta) conformations (PDB: 1DOE)[149].

A hydride is then transferred from NADPH and once reduced the isoalloxazine moves back to the *in* position. The flavin is now in a solvent free environment and reacts with O_2 to form the flavin-C4a-hydroperoxide, which then reacts with the substrate. This movement of the isoalloxazine

allows the two half reactions to occur separately from each other, reducing the likelihood of H_2O_2 formation. This movement gave rise to the term "waving flavin".

A number of other type A flavin monooxygenases have been characterised but to a lesser degree. Some of these exhibit a "waving flavin", such as phenol hydroxylase; an enzyme isolated from *Trichosporon cutaneum* that catalyses the conversion of phenol to catechol[150]. The crystal structure exhibited a major conformational change associated with the *in*/*out* conformers. The largest residue movement between the two conformers was reported as exceeding 30 Å. Other type A flavin monooxygenases, like 2-methyl-3-hydroxypyridine-5-carboxylic acid oxygenase (MHPCO), have not shown any evidence of a "waving flavin". MHPCO catalyses the oxygenation of 2-methyl-3-hydroxypyridine-5-carboxylic acid in the vitamin B_6 degradation pathway. Structural studies upon MHPCO have shown the protein only in one conformation similar to the *in* conformer[151, 152]. The presence of a solvent tunnel for substrate binding is also exhibited in these two examples.

1.5.4. Mechanism of kynurenine 3-monooxygenase

No evidence has been published that indicate KMO exhibits a "waving flavin". Despite this, overall mechanism has been proposed to be similar to that of PHBH, albeit without the "waving flavin". Using *Pf*KMO Crozier-Reabe *et al.*[142] investigated the reductive and oxidative half reactions of KMO using stopped-flow UV and fluorescence spectrophotometry. This has been used to elucidate the KMO mechanism as shown in Figure 1.15.

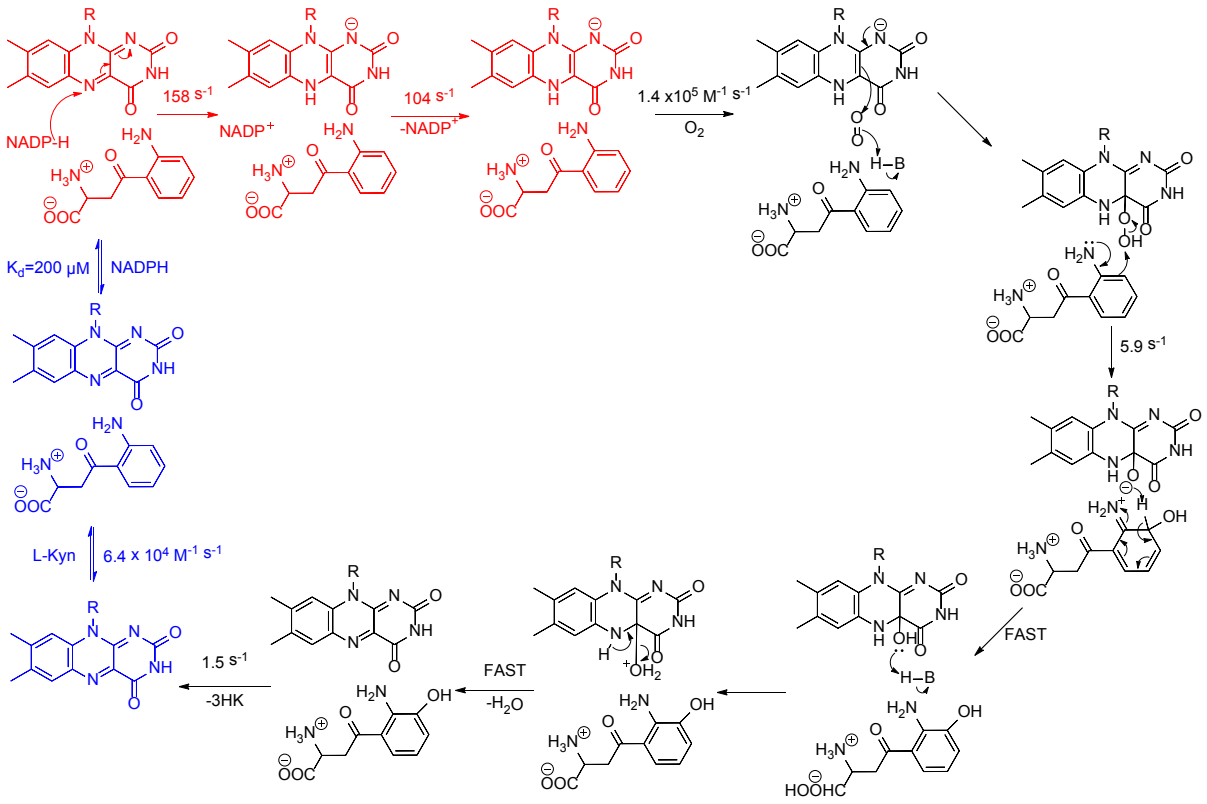

Figure 1.15 Schematic of the mechanism of *Pf*KMO. The rates of most steps are included. The mechanism is split into 3 parts: ligand binding (blue), the reductive half-reaction (red) and the oxidative half-reaction (black)[142].

As with PHBH the rates of each step in the reaction have evolved to protect against formation of H_2O_2. For example, the rate of reduction of flavin by NADPH is stimulated 10^3-fold by the presence of L-Kyn[142]. Similarly to PHBH the last step, the release of product, is the slowest and is therefore rate-limiting[142]. However, evidence showed some differences to the PHBH mechanism. For example, the release of $NADP^+$ is very rapid (104 s^{-1}) while the oxygenation of substrate is relatively slow (5.9 s^{-1}). Study of KMO also shows evidence of a single complex of [KMO(red) · L-Kyn · $NADP^+$] unlike PHBH which has two separate complexes as is required for the "waving flavin" movement, [Enzyme(ox) · Substrate · NADPH] and [Enzyme(red) · Substrate · NADP+].

1.5.4.1. *P. fluorescens* kynurenine 3-monooxygenase pharmacology

In order to assess the viability of *Pf*KMO as a surrogate for the human enzyme in a structure-based drug design approach, the pharmacology of both *Pf*KMO and hKMO has been compared[153]. The pIC50 values for 12 literature KMO inhibitors were assessed using LC/MS based analysis for *Pf*KMO and hKMO. The pIC50 values showed a correlation when compared, with the inhibitors affecting *Pf*KMO in a similar pattern to hKMO. The overall correlation had an R^2 value of 0.83 and showed that all inhibitors had a greater effect on hKMO in general. The results can be seen in Figure 1.16[153]. This finding endorses *Pf*KMO as a viable model for use in drug development for hKMO.

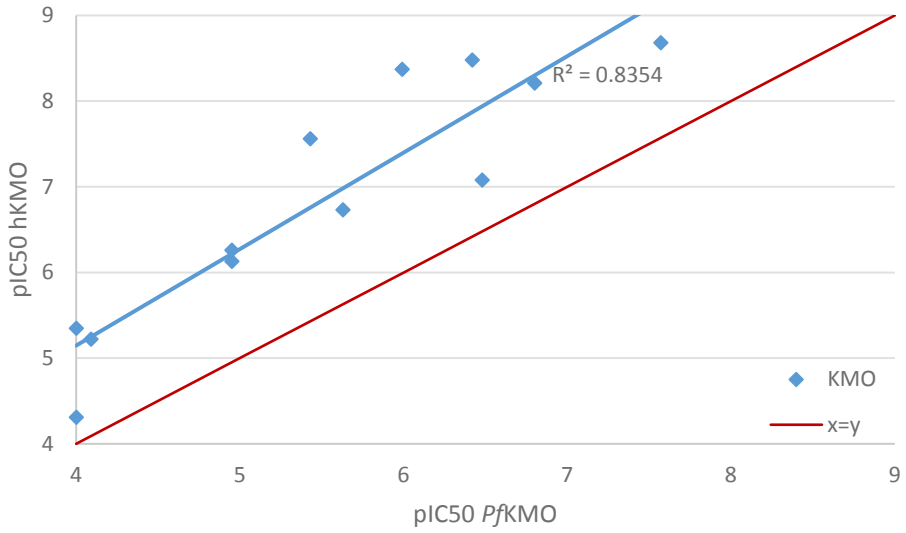

Figure 1.16 A graph to show the correlation of pIC50 values between *Pf*KMO and hKMO. Drawn using data from source[153].

The K_M of *Pf*KMO and hKMO have also been shown by Hutchinson *et al.* as comparable, 7 μM and 2 μM respectively, using a Rapidfire mass spectrometry assay[133, 154].

1.6. Structural studies of kynurenine 3-monooxygenase

1.6.1. Saccharomyces cerevisiae kynurenine 3-monooxygenase structures

The first published structures of KMO were of *Saccharomyces cerevisiae* KMO (*Sc*KMO) truncated at residue 396 or 394[140]. The 396 truncation occurred through proteolysis within the cell, while the 394 truncation was engineered. The engineered 394Δ enzyme showed turn-over of L-Kyn despite lacking the C-terminus, unlike the results obtained with pig KMO[135]. The 394 truncated enzyme was successfully crystallised both substrate-free and in complex with an inhibitor UPF 648, shown in Figure 1.17, to resolutions of 1.82 Å and 2.13 Å, respectively. The enzyme in complex with UPF 648 is shown in Figure 1.18, Figure 1.20 and Figure 1.21.

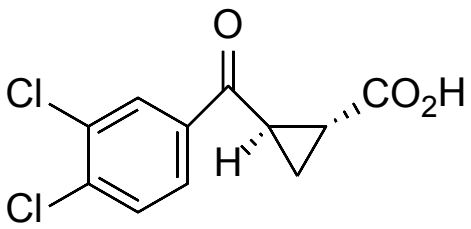

Figure 1.17 Structure of inhibitor UPF 648

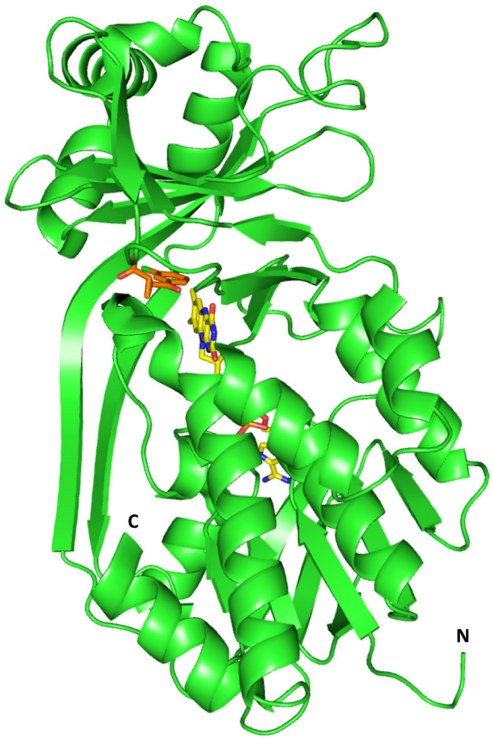

Figure 1.18 Crystal structure of *Sc*KMO bound with both FAD, shown in yellow, and inhibitor UPF 648, shown in orange (PDB: 4J36)[140]. The termini of the protein are shown here and throughout this thesis as C and N.

The structure of *Sc*KMO showed structural similarity to PHBH, reinforcing the position of PHBH as a model enzyme for type A flavoprotein aromatic hydroxylases. When overlaid using root mean square deviation (RMSD) of alpha carbons the alignment showed a RMSD value of 10.6 Å. As shown in Figure 1.19, the FAD molecules take up similar positions within the proteins. See section 2.8.2 for further details on alignments.

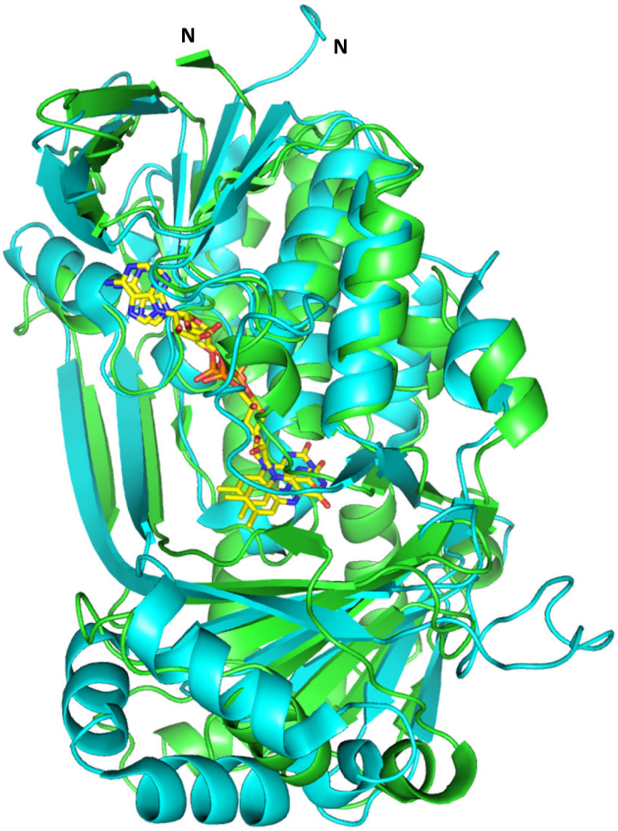

Figure 1.19 Overlay of the crystal structures of *Sc*KMO in cyan (PDB: 4J36) and *Sc*PHBH in green (PDB: 1PHH) both bound with FAD, shown in yellow[140, 155]. RMSD = 10.6 Å from 291 aligned Cα atoms.

Some of the key binding residues interacting with the UPF 648 carboxylate were R83 and Y97; these are conserved throughout the KMOs, see Figure 1.10[140]. The aromatic part of the inhibitor was surrounded by hydrophobic residues, many of which are also conserved. The importance of residue R83 was demonstrated through mutagenesis[140]. When substituted by alanine or methionine the enzyme exhibited 25% and <3% of wild-type activity, respectively. The substitutions also led to a 20-fold drop in binding affinity for UPF 648.

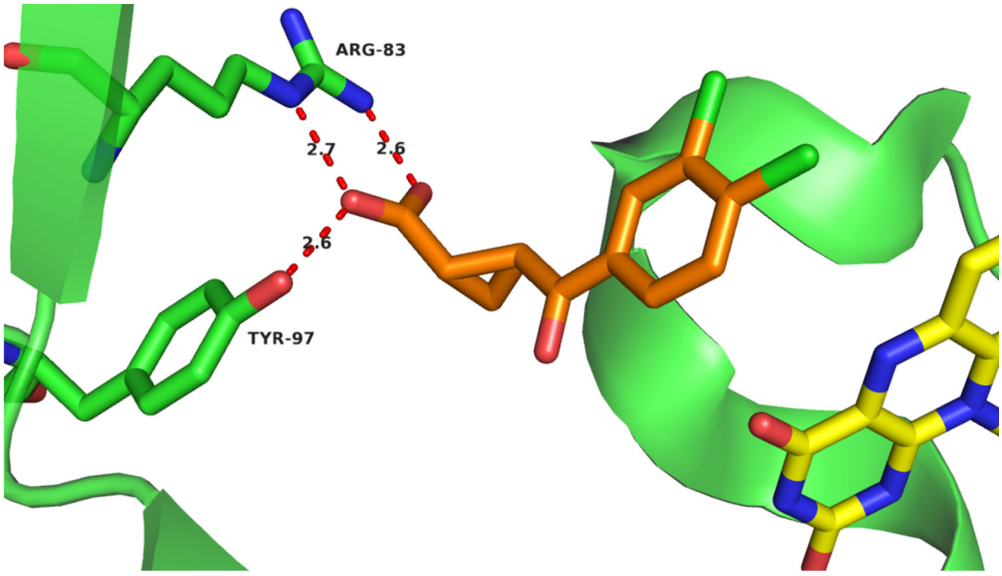

Figure 1.20 Crystal structure of the active site of *Sc*KMO showing interactions between UPF 648 and residues R83 and Y97 (PDB: 4J36)[140]. Distances between atoms are in Å and shown as red dashed lines.

A change in structure is seen when *Sc*KMO binds UPF 648; the loop P321-Q325 re-orientates increasing the disorder of the C-terminus α-helix. This movement is caused by F322 moving away from the active site to accommodate the two chloride groups in UPF 648 that have no equivalent groups in L-Kyn. These conformational changes can be seen in Figure 1.21.

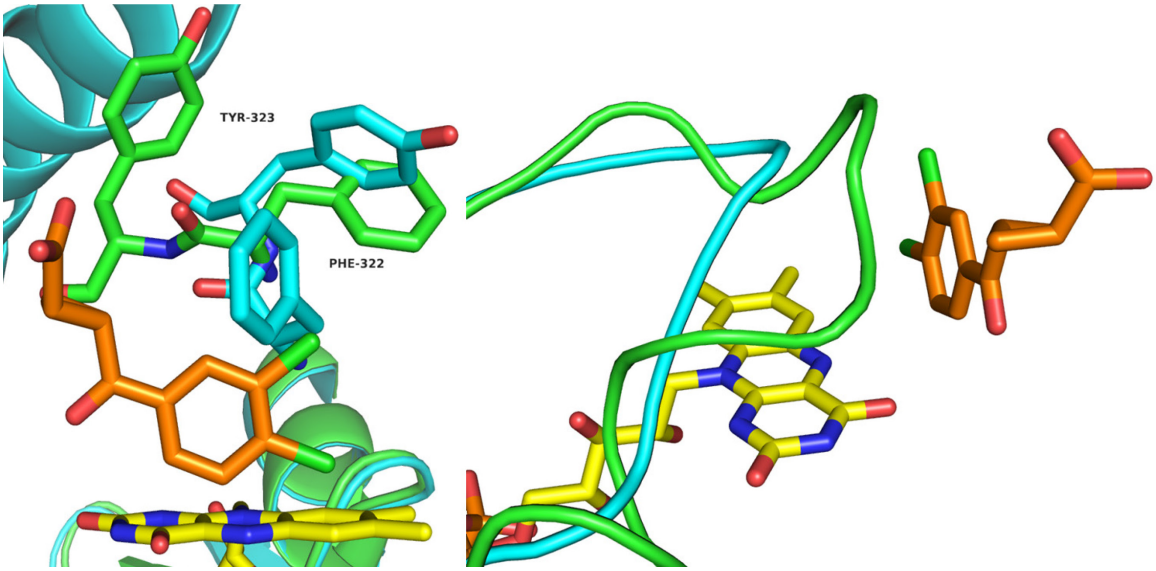

Figure 1.21 Crystal structure showing the movement exhibited by *Sc*KMO upon UPF 648 binding. The UPF 648 bound model is shown in green, the unbound model shown in blue. RMSD = 0.793 Å from 349 aligned Cα atoms. Left shows residues F322 and Y323, right shows the loop P321-Q325 (PDB: 4J36, 4J33)[140].

Despite advances in understanding of active site interactions some potentially important residues were impossible to investigate due to the C-terminus truncation of *Sc*KMO. The truncation also precluded the investigation of any conformational change within the C-terminus upon ligand binding.

1.6.2. Pseudomonas fluorescens kynurenine 3-monooxygenase structures

In early 2016 a structure of *Pf*KMO in complex with inhibitor GSK180 was published to a resolution of 3.19 Å, see Figure 1.22. The importance of residue R84, the equivalent of *Sc*KMO's R83, was demonstrated by the existence of an apparent salt bridge between it and the carboxylate group of the inhibitor. The residues Y98 and N369 also form H-bonds with the carboxylate, while Y404 is shown interacting with the carbonyl.

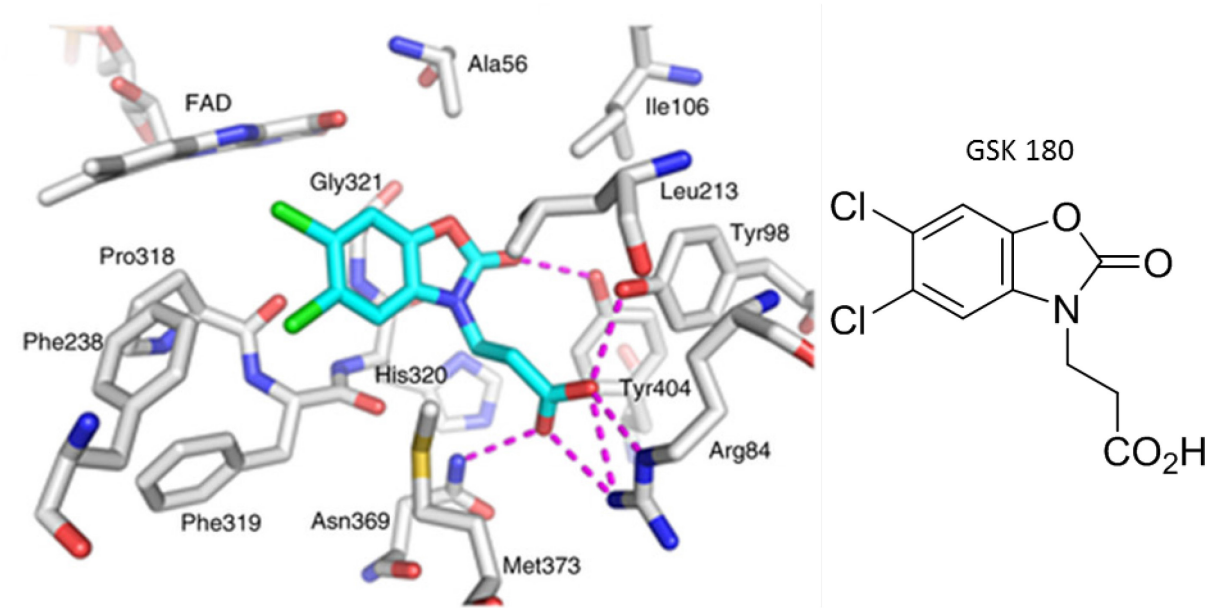

Figure 1.22 Left shows the crystal structure of the active site of *Pf*KMO interacting with the inhibitor GSK180, the structure of which is shown on the right. Image adapted from Mole *et al.* (PDB: 5FN0) [11].

Structural study of *Pf*KMO has been on going within the Mowat group for several years. Crystals were originally grown in conditions using HEPES and polyethylene glycol 4000 (PEG 4K) but did not diffract well. The protein itself was inherently unstable and lost activity over time in the absence of a reductant. It was hypothesised that this was due to disulfide bonds forming between molecules[141]. Two cysteine to serine mutations were introduced at the C-terminal residues to give C252S/C461S, a double mutant referred to in this work as dm2[143]. This greatly improved the stability of the protein, as shown in Table 1.2. The kinetic analysis showed there was no significant difference in turnover or substrate binding between dm2 and the wild type *Pf*KMO. All mutants and kinetic data were produced by Martin Wilkinson[143].

Protein	K_M (μM)	k_{cat} (s^{-1})	$t_{1/2}$ (hours)
Wild type	14.5 ± 2.5	9.1 ± 1.6	7.9 ± 2.6
C252S	18.7 ± 5.9	6.8 ± 0.6	53.5 ± 24.3
C461S	15.5 ± 4.2	6.4 ± 0.5	31.5 ± 2.3
dm2	8.8 ± 1.5	8.9 ± 1.1	75.2 ± 5.2

Table 1.2 Michaelis-Menten kinetics and enzyme half-life for *Pf*KMO and cysteine to serine mutants. $t_{1/2}$ calculated from time required for half enzyme activity to be lost when incubated at 25 °C[143].

Crystals of dm2 diffracted to a greater degree and using selenomethionine substitution and multi-wavelength anomalous dispersion a structure of substrate bound dm2 was solved. The space group was $I4_122$ with a solvent content of over 80% leading to a relatively low resolution of 3.4 Å. This led to the first crystal structure of KMO with an intact C-terminus and the first in complex with L-Kyn, this is shown in Figure 1.23.

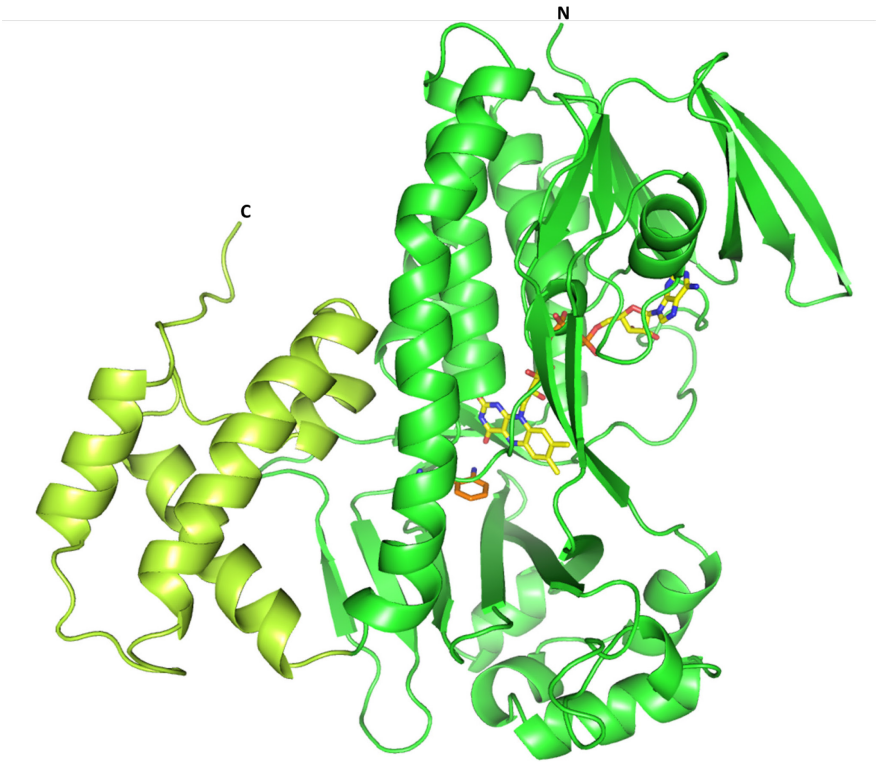

Figure 1.23 Crystal structure of dm2 bound with L-Kyn (orange)[143]. The C-terminal domain is shaded light green.

Despite the low resolution the presence of substrate was confirmed and the positioning of the substrate molecule could still be determined, as seen in Figure 1.24. Interactions with key binding residues could also be identified and these can be seen in Figure 1.25. The residues identified matched with equivalents in the previously mentioned *Sc*KMO work, including R84 and Y98. Other residues, including those

impossible to characterise in the truncated *Sc*KMO, were recognised as having substrate interactions, such as N369, Y404 and H320.

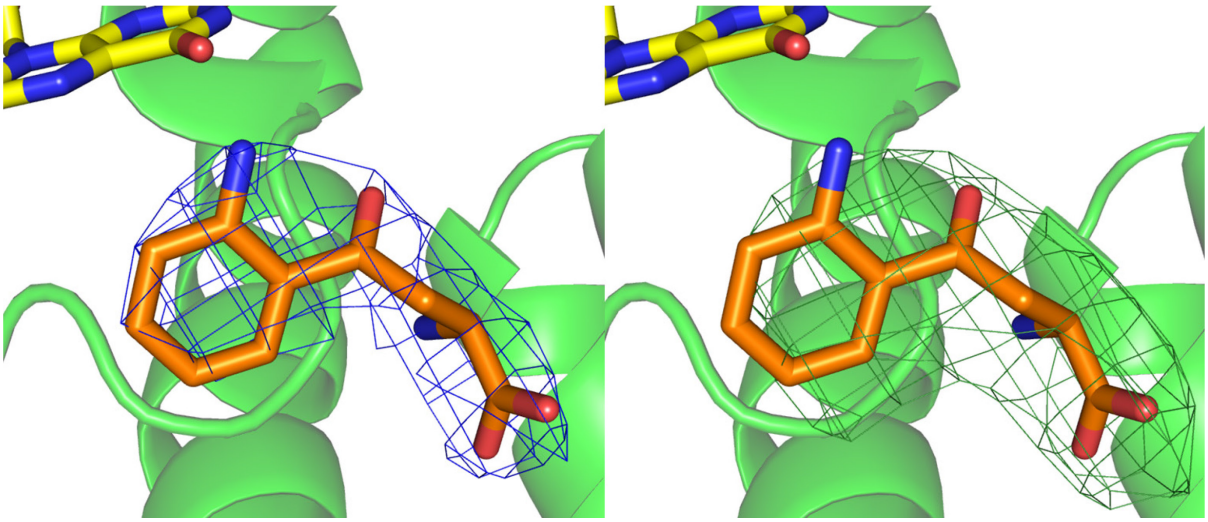

Figure 1.24 Crystal structure of dm2 bound with L-Kyn[143]. The 2Fo-Fc map is shown in blue (contoured at 0.12 eÅ^{-3}, 1.50 σ) and the Fo-Fc map, calculated in absence of L-Kyn, is shown in green (contoured at 0.10 eÅ^{-3}, 3.0 σ).

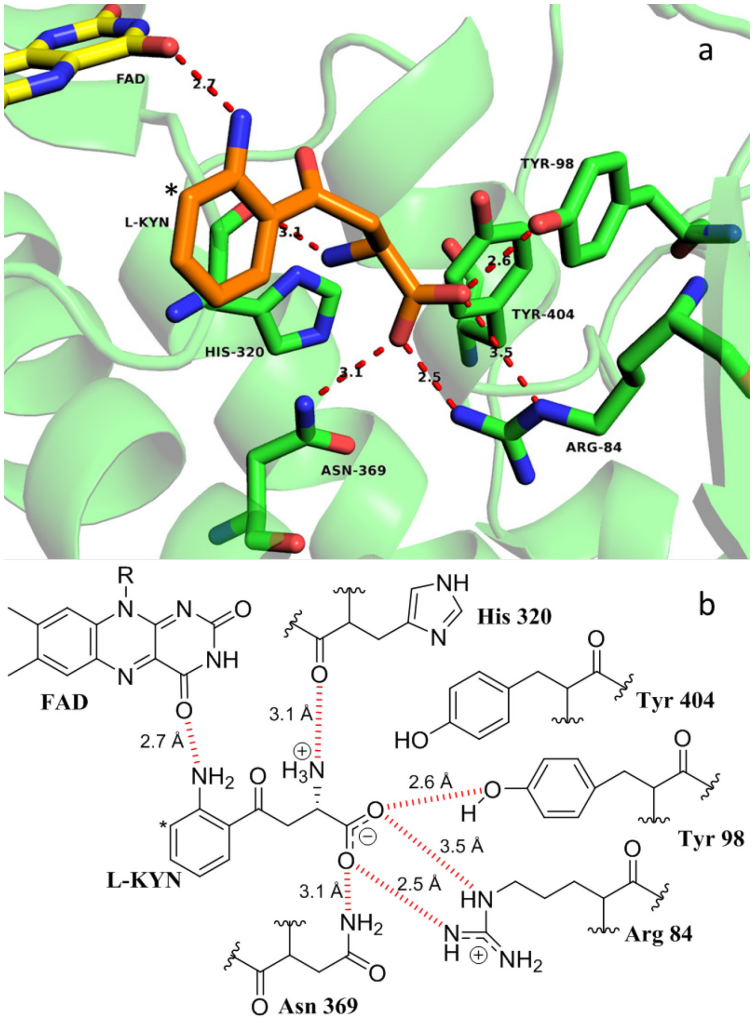

Figure 1.25 a. Crystal structure of the active site of dm2 bound with L-Kyn (orange). **b.** Chemdraw schematic of dm2 bound with L-Kyn. Distances between atoms are in Å and shown as red dashed lines. The site of oxidation is denoted with an asterisk (*).

1.6.3. C-terminus movement

The structure of ligand-free dm2 was also solved to a resolution of 2.26 Å[143]. Crystals belonged to space group $P2_122_1$ and showed a solvent content of roughly 50%. The majority of the protein showed high similarity with the substrate-bound model. There was, however, a significant movement observed within the C-terminus between the bound and unbound structures. Compared to the substrate-free form, the substrate-bound form showed a movement of up to 6.1 Å at residue 381 and implied the existence of *open* and *closed* conformations. As seen in Figure 1.26, the surface of the protein shows a clear channel to the active site in the *open* position, which then disappears when in the *closed* position. This appears similar to the conformational change seen in PHBH, as discussed in section 1.5.3.2, which serves to exclude H_2O from the active site.

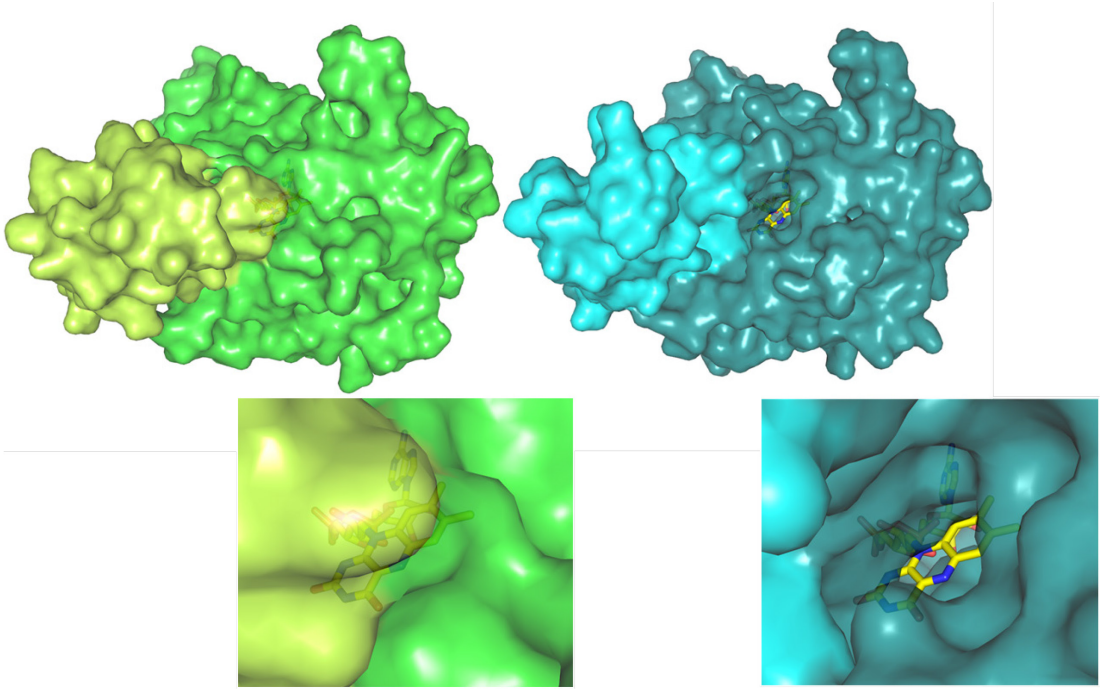

Figure 1.26 Surfaces of crystal structures of dm2. The substrate bound *closed* position is shown on the left, the substrate free *open* position, with channel leading to isoalloxazine of FAD, is shown on the right. The C-terminus is shown in a lighter shade. An enlarged view of the ligands is also shown.

The positioning of residues between the mobile C-terminus and the main body of the protein were investigated and a putative network of interactions was identified. In the substrate-bound model many of the residues are within electrostatic or H-bonding distance of each other and also of L-Kyn, with R84 playing a key role between the two. In the substrate-free model the residues are well out of range for any interaction, as seen in Figure 1.27.

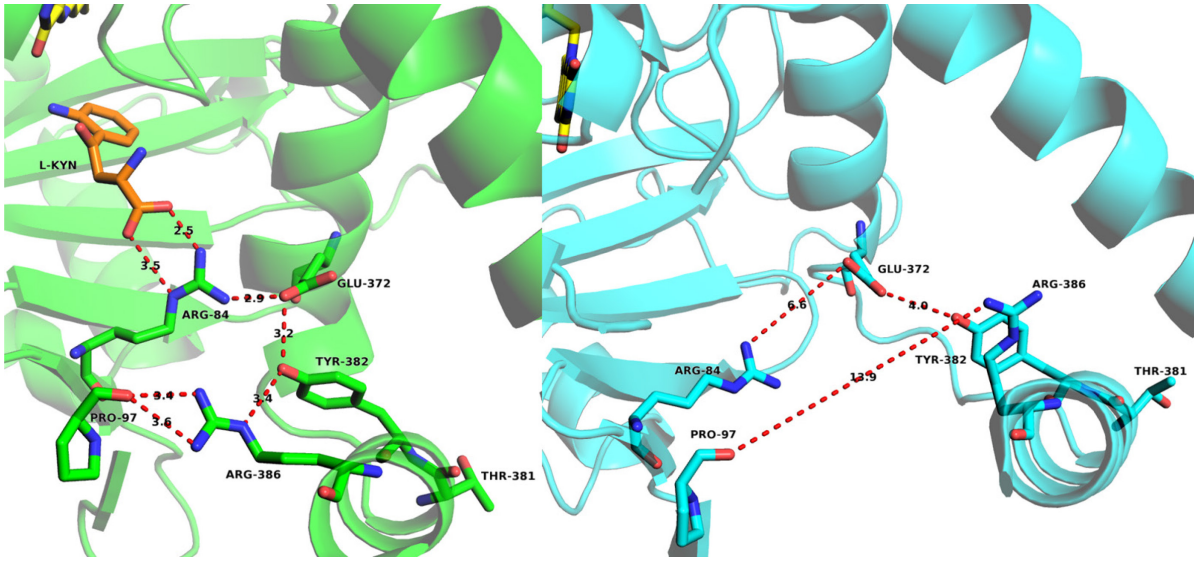

Figure 1.27 Crystal structures of dm2, showing distances between residues in the putative bonding network of the C-terminus[143]. Left shows the *closed* position, right shows the *open* position. Distances between atoms are in Å and shown as red dashed lines.

This proposed network was investigated further by site-directed mutagenesis of the key residues involved, and analysis of the crystallography and kinetics of these mutant forms of PfKMO.

A comparison of C-terminus swing was made based on the movement of the C_α atom of T381, comparing the original substrate-bound dm2 structure to the structure in question[143]. This is shown in Table 1.3. A decrease in C-terminus movement was seen in all mutants, placing the substrate-bound C-terminus in an intermediate position between *open* and *closed* conformations, known as a *mixed* position.

Protein	K_M (μM)	k_{cat} (s^{-1})	C-terminus swing (Å) bound/unbound
dm2	8.8 ± 1.5	8.9 ± 1.1	0.0/6.1
E372T dm2*	10.3 ± 0.7	8.7 ± 0.1	3.7/n.a.
R386K dm2	17.0 ± 1.4	12.0 ± 0.9	4.2/6.1
R386T dm2	63.9 ± 2.4	15.2 ± 1.3	4.3/6.1
Y382F dm2	17.1 ± 1.0	12.7 ± 0.2	3.8/n.a.

Table 1.3 Summary of Michaelis-Menten kinetic data and extent of C-terminus swing for mutant forms of dm2[143]. (n.a. is shown where no data exist). *E372T mutant also contained a L367R mutation.

The space groups of each substrate-bound mutant structure were also $I4_122$ leading to a high solvent content and low resolution, the highest being 3.24 Å. The fact that each mutant showed the same space group, even with a more open conformation, implied that the C-terminus movement was dependent on the protein and not the crystal packing. This does not, however, totally remove the possibility that the

presence of L-Kyn is responsible for the change in space group. The different space groups could in turn form different crystal contacts that could be the cause of the C-terminus movement.

A third space group, $P6_122$, was observed in crystals of dm2 bound to a number of tightly binding inhibitors based upon the known KMO inhibitor Ro 61-8048[84]. These inhibitors showed K_i values between 0.02 and 2.77 μM. They bound to dm2 much like L-Kyn, but all the molecules within the ASU took up a conformation similar to the *open* unbound form. This was rationalised as the tails of the inhibitors disrupting the interactive network of the C-terminus[143].

Overall the hypothesis of C-terminus dynamism being linked to mechanism still requires further confirmation before firm conclusions can be drawn.

The resolution of *Pf*KMO structures allowed for identification of potentially important residues. They showed possible interactions in the active site and C-terminus but did not allow for detailed analysis of these interactions. Due to co-ordinate error, in a 3.4 Å resolution structure with high overall B-factors, a distance measured as 3.2 Å associated with a weak H-bond may in-fact be a close strong interaction or a very weak distant interaction.

1.6.4. Kinetic analysis of dm2 mutants

The residues highlighted within the active site, including R84, Y404, H320, and N369 were assessed for their importance in enzyme activity by site-directed mutagenesis. The conservative mutations were Y404F, R84K, N369S and H320F. The mutants were then characterised via Michaelis-Menten kinetic analysis. All mutants and kinetic data were produced by Helen Bell[156].

Enzyme	K_M (μM)	k_{cat} (s^{-1})	k_{cat}/K_M (s^{-1} μM^{-1})
dm2	8.8 ± 1.5	8.9 ± 1.1	1.01
Y404F	96.5 ± 4.0	6.2 ± 0.1	0.06
R84K	0.0	0.0	0.0
N369S	27.7 ± 3.3	4.5 ± 0.2	0.16
H320F	18.0 ± 1.2	7.1 ± 0.2	0.39

Table 1.4 Summary of Michaelis-Menten kinetic data for mutant forms of dm2. R84K showed no measurable activity and is reported as zero[156].

The variants that showed the greatest change in turnover and binding were Y404F and R84K. The efficiency of Y404F is decreased by a factor of ~10 and the R84K enzyme showed no turnover whatsoever, totally inactivating the enzyme. These residues were hence identified as the most interesting in terms of their influence on function.

Crystal structures of these mutants were solved but suffered once again from low resolution and high solvent content, especially when substrate was bound. This made it difficult to assign a position to L-Kyn within the active site.

1.7. Aim

Throughout this chapter it can be seen that KMO is a viable drug-target for a number of diseases including Huntington's disease, Alzheimer's disease, brain ischemia and acute pancreatitis. *Pf*KMO dm2 has been used as a model for hKMO and has been involved in the development of a lead molecule. In an effort to validate previous structural models, further characterisation of key binding residues is therefore necessary. Better understanding of the mechanism of KMO is also required, particularly concerning the role of the C-terminus and the static nature of the FAD.

Despite the information previously gleaned from both kinetic and crystallographic work the theories surrounding *Pf*KMO require further clarification. The previous crystal structures of *Pf*KMO highlighted potentially important residues within the active site and a putative C-terminus bonding network. However due to low resolution data the exact positions and roles of these residues are yet to be pinpointed. The kinetic data collected on *Pf*KMO mutants has verified the importance of certain residues within the active site, but the reasoning behind these changes in kinetic data is yet to be fully established.

The aim of this work is to rationalise the importance and function of these highlighted residues in the active site and to clarify the presence of a C-terminus H-bonding network within dm2 through high-resolution structural data.

Chapter 2: Materials and Methods

2.1. Materials

All water was purified using a Sartorius arium 61316 water purifier. All chemicals used were ≥97 % pure. All chemicals were purchased from Sigma Aldrich unless otherwise stated.

2.2. DNA cloning

2.2.1. Primer sequence design

Primers used in Sanger sequencing are listed in Table 2.1. Each was purchased in dry form, dissolved in water to a concentration of 100 μM and stored at -20 °C.

Primer	Sequence
T7 promoter	TAATACGACTCACTATAGGG
T7 terminator	GCTAGTTATTGCTCAGCGG

Table 2.1 Primer sequences used in Sanger sequencing.

2.2.2. pET plasmid transformation

pET plasmids containing the gene of interest were used to transform BL21 (DE3) *E. cloni*® EXPRESS chemically competent cells sourced from Lucigen®. Cells were removed from -80 °C storage and thawed on ice, and 40 μL of cells was transferred to a pre-chilled sterile culture tube (15 mL). 1 μL DNA was added and briefly stirred with a pipette tip. The cells were incubated on ice for 30 mins, heat shocked in a 42 °C water bath for 45 seconds and returned to ice for 2 minutes. 960 μL Lucigen® Expression Recovery Medium was added to the tube and incubated for 1 hour at 37 °C with shaking at 250 rpm. 100 μL of cells was plated onto LB agar plates (see section 2.3.1) containing 50 μL/g ampicillin and incubated overnight at 37 °C.

2.2.3. Plasmid extraction

10 mL LB with 50 μg/L of the relevant antibiotic was inoculated with cells in 30 mL sterile tubes and incubated at 37 °C with shaking at 200 rpm for 16 hours. The cells were harvested by centrifugation for 10 minutes at 8000 rpm at 4 °C in 60 mL flasks using a Sorvall RC 6 Plus centrifuge and a Sorvall SS-34 rotor. The plasmid was extracted from the cell pellet using a QIAprep Spin Miniprep Kit and eluted with 40 μL elution buffer into a sterile Eppendorf tube.

2.2.4. Cell stock preparation

900 μL of sterile LB media (see section 2.3.1) with no antibiotic was placed in a sterile Eppendorf tube. Cells were introduced to the media; the tube inverted several times and incubated at room temperature for 10 minutes. 100 μL of dimethyl sulfoxide (DMSO) was added and the tube stored at -80 °C for long-term storage.

2.2.5. Sanger sequencing

Sanger sequencing was performed by the University of Edinburgh based company The Genepool using BigDye v3.1 Terminator Cycle Sequencing Kit from Thermofisher. Each sample was prepared in a 0.2 mL strip tube with 5 μL plasmid DNA and 1 μL sequencing primer (3.2 μM). Each plasmid was sequenced in both forward and reverse using primers for T7 promoter or T7 terminator respectively. The results were combined to show the full gene sequence.

2.3. Cell growth and Harvest

All protein-producing bacteria were recombinant BL21 *E. coli* containing the λDE3 lysogen that carries the T7 RNA polymerase gene.

2.3.1. Media

All bacteria were grown in Luria Bertani media. To prepare Luria Bertani (LB) media the following were dissolved in water:

10 g/L tryptone (Sigma Aldrich)

5 g/L yeast extract (Scientific Laboratory Supplies)

5 g/L NaCl

All media were adjusted to pH 7.0 with NaOH and autoclaved at 121 °C for 20 minutes.

All subsequent supplements to the media, such as ampicillin, were sterile filtered through 0.45 μm Millipore filters.

LB agar plates were prepared by adding 15 g/L agar to LB media before autoclaving. The solution was cooled to 40 °C before antibiotic was added, the media poured into sterile Petri dishes and allowed to cool.

To prepare X-Gal/IPTG plates 10 μL X-Gal solution (20 mg/mL) and 10 μL isopropyl β-D-1-thiogalactopyranoside (IPTG) (100 mM) per 1 mL of media was added at the same time as the antibiotic.

2.3.2. *Pf*KMO growth

All *Pf*KMO proteins were produced using a pET17b expression vector containing the appropriate KMO gene. 10 mL aliquots of LB with 50 µg/L ampicillin were inoculated with cells in 30 mL sterile tubes and incubated for 24 hours at 37 °C with shaking at 180 rpm. These 10 mL growths were then added to separate 1 L baffled flasks containing 1 L LB media supplemented with 50 µg/L ampicillin and incubated at 22 °C with shaking at 180 rpm for 16 hours.

2.3.3. Cell Harvest

Cells were harvested by centrifugation for 10 minutes at 2704 x *g* at 4 °C in 500 ml flasks using a Sorvall RC 6 Plus centrifuge and a Sorvall SLA-3000 rotor. The cell pellets were stored at -20 °C until required.

2.4. Protein purification

2.4.1. *Pf*KMO purification

The buffers used in *Pf*KMO purification are listed in Table 2.2. All buffers were pH adjusted using NaOH or HCl and filtered through a 0.45 µm filter.

Lysis buffer	20 mM HEPES pH 7.5 10 mM NaCl 1 mM DTT
Wash buffer	20 mM HEPES pH 7.5 50 mM NaCl 1 mM DTT
Elution buffer	20 mM HEPES pH 7.5 110 mM NaCl 1 mM DTT
Ammonium sulfate solution	20 mM HEPES pH 7.5 75 % saturated ammonium sulfate 1 mM DTT
Size exclusion buffer	20 mM HEPES pH 6.8 10 mM NaCl 1 mM DTT

Table 2.2 Compositions of *Pf*KMO purification buffers.

Cell pellets were thawed on ice and resuspended in 10 mL lysis buffer per gram of cell mass. The suspension was supplemented with phenylmethanesulfonyl fluoride (PMSF) to a concentration of 0.2 mM. The solution was ultrasonicated on ice in 20-second bursts for a total of 2 minutes with cooling between bursts. This suspension was then centrifuged for 1 hour at 47808 x *g* at 4 °C in 60 mL tubes using a Sorvall RC 6 Plus centrifuge and a Sorvall SS-34 rotor.

The supernatant was filtered through 0.45 μm syringe filters and loaded onto a 60 mL FPLC Q Sepharose 26/10 column pre-equilibrated with lysis buffer. 400 mL wash buffer was then used to wash the column. The protein was then eluted using 500 mL elution buffer. The protein was collected by an automated fraction collector. The fractions were tested for *Pf*KMO content using an absorbance of at 450 nm and for activity using a kinetic assay described in section 2.6.1.

In order to precipitate the protein, pooled fractions of *Pf*KMO were mixed slowly on ice with 75% saturated ammonium sulfate solution to reach a final saturated concentration of 50%. After 30 minutes incubation on ice the suspension was centrifuged for 20 minutes at 47808 x *g* at 4 °C in 60 mL tubes using a Sorvall RC 6 Plus centrifuge and a Sorvall SS-34 rotor. The pellets could be stored at -20 °C prior to further purification.

The pellets were resuspended on ice using minimal amounts of size exclusion buffer. The solution was concentrated by centrifugation using VIVASPIN MW 30000 concentrators to a volume of 2 mL. This was loaded onto a 320 mL 26/600 Superdex 75 column pre-equilibrated with size exclusion buffer. The protein was collected by an automated fraction collector monitored by UV absorbance at 280 nm.

The fractions were tested for purity and concentration using absorbance measurements described in section 2.5.2 and pooled. The concentration of protein was adjusted as necessary for future use and the protein was aliquoted and flash frozen in liquid nitrogen before being stored at -80 °C. The yield of each preparation varied between 1-2 mg of pure protein per litre of LB media in the initial growth.

2.5. Protein purity analysis

2.5.1. SDS-PAGE analysis

Liquid samples were mixed in a 1:2 ratio with novex® NuPAGE® LDS sample buffer (4×), small scrapings of solid samples were mixed with 20-40 μL of the same buffer to prevent overloading. Samples were boiled for 5 minutes and loaded onto a pre-cast novex® NuPAGE® 4-12% Bis-Tris gel in novex® NuPAGE® MES SDS running buffer. 4 μL novex® SeeBlue® Pre-Stained Standard was added as a protein marker. Gels were run at 150 V for 1 hour. Gels were stained for 1 hour with gentle shaking using a solution comprising 0.1% Coomassie Brilliant Blue G-250, 40 % methanol, 10 % acetic acid and 50% water. The gel was washed and de-stained with water overnight.

2.5.2. UV-Vis spectrophotometry

All UV-Vis spectrophotometry was performed using a Cary UV-Vis 50 Probe spectrophotometer with a quartz cuvette of path length 1 cm. To assess concentration and purity of protein solutions a scan between 600-240 nm was performed using the relevant buffer as a blank.

The concentration was calculated using the Beer-Lambert law using the molar absorbance for flavin at 450 nm (12300 M^{-1} cm^{-1}). Using prior experience purity was analysed by comparing the aromatic residues absorbance at 280 nm to the flavin absorbance at 450 nm with a respective ratio of 7:1 considered pure.

2.5.3. Mass Spectrometry

Any liquid chromatography electrospray ionisation mass spectrometry was performed using 15 µL of ~15 µM protein. This was loaded onto an Aeris Widepore 3.6 µm C4 50 x 2.1 mm column connected to a Waters ZSpray Lockspray electrospray ionisation source. This was then analysed using a Synapt™ G2 time of flight mass spectrometer. The peaks analysed were those that gave a protein envelope spectrum. The program used to visualise and calculate the mass of the peak was *Masslynx* (Waters, Version V4.0).

2.6. Protein Kinetic Analysis

2.6.1. *Pf*KMO Kinetic assay

Kinetic activity was monitored using a Cary UV-Vis 50 Probe spectrophotometer monitoring at 340 nm over 30 seconds with a quartz cuvette of path length 1 cm. This absorbance corresponds to the peak for NADPH, and this falls as it is consumed (to form $NADP^+$) as it reduces *Pf*KMO. All data were collected at 25 °C controlled by a thermostatic water bath. The assay buffer components and final concentrations are shown in Table 2.3.

Assay component	Final Concentration
HEPES buffer, pH 8.0	20 mM
NADPH	220 µM
L-kynurenine	Various
*Pf*KMO	0.1-0.5 µM

Table 2.3 Composition of *Pf*KMO kinetic assay.

Results were given in change in absorbance per minute; these were converted to enzyme turnover using Equation 2.1.

$$Turnover = \frac{\Delta\ Absorbance\ (340\ nm)}{60\ \times\ \varepsilon\ (NADPH\ at\ 340\ nm) \times\ [enzyme]}$$

Equation 2.1 Calculation of enzyme turnover (s^{-1}) where Δ Absorbance is the change in absorbance at 340 nm (min^{-1}), ε is the extinction coefficient of NADPH at 340 nm (6200 M^{-1} cm^{-1})[157] and [enzyme] is the concentration of enzyme (M).

2.6.2. Data modelling

In order to determine K_M and k_{cat} kinetic data were fitted using the Michaelis-Menten function in *Origin 2015* (OriginLab, Northampton, MA).

2.7. Crystallography

2.7.1. Hanging-drop vapour diffusion

Linbro 24-well plates were used to grow crystals. Each well was filled with 0.5 mL well solution. Well solutions varied in concentration of polyethylene glycol 8000 (PEG 8K) and pH to increase likelihood of crystal growth. The components and concentrations of well solution are shown in Table 2.4.

Well solution component	Final Concentration
PEG 8K	11-21%
Sodium cacodylate pH 5.5-7.0	0.08 M
Calcium acetate	0.16 M
Glycerol	20%

Table 2.4 Composition of crystal well solution.

1 µL of purified ~250 µM *Pf*KMO was placed on a siliconized glass cover slip and mixed with 1 µL of well solution. The slide was sealed onto the well using paraffin grease and the well plate wrapped in aluminium foil to block out light and stored at 18 °C in a vibration free room.

2.7.2. Streak seeding

To encourage crystal growth streak seeding was sometimes employed. A crystal of the relevant protein was crushed in the hanging drop using a glass crystal crusher. 5 µL of well solution supplemented with PEG 8K to a concentration of 25 % was added to the drop. This was then placed in an Eppendorf tube on ice, to which 45 µL well solution was added and the tube vortexed for 1 minute. Serial dilutions of the seeding solution were made using mother liquor supplemented with PEG 8K to a concentration of 25 %; each solution was vortexed prior to further dilution. These solutions were stored at -80 °C as required.

To seed the hanging drop a seeding tool (synthetic hair) was dipped into a seeding solution and dragged horizontally through the droplet, the well was then sealed as normal.

2.7.3. Co-crystallisation and crystal soaking

In order to introduce substrate into the active site, samples were co-crystallised with L-kynurenine, or crystals were soaked in L-kynurenine solutions. Co-crystallisation was achieved by mixing pure protein sample with L-kynurenine solution to give a final concentration of 1 mM L-kynurenine prior to crystallisation.

To soak crystals they were extracted from the hanging drop and soaked in a 1 μL drop of mother liquor stabilised with 25 % PEG 8K and supplemented with L-kynurenine (1 mM). The higher PEG 8K concentration ensured crystals did not dissolve. These were then returned to the sealed well for at least 12 hours to allow substrate to enter the active site.

2.7.4. Crystal mounting

Crystals were extracted from their droplet using CrystalCap™ SPINE HT 0.1-0.2 mm Cryoloops. Where necessary crystals were cryogenically protected by dipping them sequentially in well solution containing 10 % ethylene glycol, then 20 % ethylene glycol. The crystals were flash cooled in liquid nitrogen, placed in a sample puck and shipped to Diamond Light Source synchrotron.

2.7.5. Data collection

Data were collected remotely from Diamond Light Source synchrotron. Crystals were screened by taking 4 images 90 ° apart, strategies were calculated using *EDNA* or *Mosflm*. The typical data collection parameters of a data set are shown in Table 2.5.

Oscillation	0.05-0.15 °
Transmission	20-30%
Exposure	0.05-0.2 seconds
Resolution	1.2-2.0 Å
Total oscillation	180-360 °

Table 2.5 Synchrotron collection parameters.

2.8. Data Processing

2.8.1. Model building

Data processing was performed using the *CCP4* program suite[158]. Images collected from the synchrotron were indexed, integrated, scaled and merged either by using *iMosflm* and *Scala* or using *Xia2*[159-161]. In all cases 5 % of reflections were used to make a FreeR test set.

Models were created from the resulting data using the molecular replacement program *Molrep*[162]. The original model used for molecular replacement in this thesis was 1 molecule of substrate bound dm2 in space group $I4_122$ solved by Martin Wilkinson of the Mowat group[143, 162]. Water molecules were removed from the model before use, FAD and L-Kyn remained. *Molrep* employed non-crystallographic symmetry (NCS) of the original $I4_122$ model to locate the second molecule within the asymmetric unit (ASU).

Restrained refinement was carried out using *Refmac5*, alterations to models were made using *Coot*[163, 164]. B-factors were modelled isotropically in all cases. Further calculations involving B-factors, such as averaging, was performed using all atoms. Neither NCS nor translation-libration-screw (TLS) modelling were used throughout refinement. Simple Wilson scaling was used for bulk solvent regions. FAD was inserted into each model using the *Coot* (version 0.7.2.1) library after the first refinement. Models were altered to decrease peaks in the DELFWT map and to adhere to structural considerations e.g., Ramachandran angles. Ramachandran angles were analysed using the *Dynarama* plug in of *Coot*. The underlying distributions of Ramachandran angles were taken from the Richardson 500[165]. Water molecules were introduced using *Refmac5* where the DELFWT map was greater than 2 σ. These waters were checked for proximity to other atoms and their effect on the DELFWT map. Further water molecules were introduced manually using peaks in the DELFWT map as a guide.

In L-Kyn exposed models, water molecules within the active site were removed to allow substrate to be modelled. L-Kyn was inserted into models using the *Coot* (version 0.7.2.1) library. L-Kyn was then manipulated to best fit the electron density within the active site. The torsion angles involving the planar ring allowed no deviation, the butanoic acid tail was allowed some variation with estimated standard deviation of 20°.

Each dataset was evaluated on a number of factors. These factors included a high correlation coefficient ($CC_{1/2}$), low R_{merge}, a high $I/\sigma(I)$, high completeness and high multiplicity. These were considered alongside the resolution of each structure and the quality of electron density, particularly around the area of interest. Only the best examples from the many datasets collected for each protein are presented in this work.

2.8.2. Model Analysis

Pymol was used to create images of the models[166]. *Pymol* was also used to create alignments of protein molecules[167]. The molecules were first aligned by residue using BLOSUM62 weightings[168]. The alpha carbon atoms (Cα) of aligned residues were matched, the molecules were superpositioned by minimising the root mean square deviation (RMSD) between matched atoms. The RMSD and number of Cα atoms aligned is recorded with each alignment. *Fft* was used to create Fo-Fc and 2Fo-Fc maps to display in images of the models[158].

In order to assess the impact of crystal contacts upon the movement of the C-terminus the program *NCONT* was used[169]. The program was limited to only analyse from residue 370 to the C-terminal residue. The calculations included the other molecule within the ASU and any symmetry generated molecules. Any interactions within the molecule itself were discounted. The contacts reported were within 5.0 Å.

Polder maps were employed to assess the likelihood of substrate occupancy using the program *PHENIX*[170]. An OMIT map is when a molecule, such as the substrate, is purposefully removed from a model. This model is then refined to create an Fo-Fc map that may show positive peaks in the region of the omitted molecule. This indicates the absence of electron density does not match the data. These simple OMIT maps do not take into account the contribution to electron density from the bulk solvent modelled across the structure. Polder maps are similar to OMIT maps but they remove this bulk solvent from the desired region and are therefore less biased by the solvent[171].

In order to validate these maps the correlation between three different polder maps are assessed. Map 1 and 2 are calculated while map 3 is the actual experimental data. Map 1 is made assuming the omitted atoms are present; map 2 is made assuming the omitted atoms are not present. If map 3 shows the omitted atoms then it will correlate best to map 1. If map 3 shows only solvent then it will correlate best to map 2.

Surface calculations were performed with the *Pymol* plugin *KVfinder*[172]. The surface calculations were to analyse the surface area and void volume of the active site of KMO, including the opening discussed in section 1.6.3. The calculations were based on the positioning of the L-Kyn molecule within the model of the *closed* molecule of substrate-bound dm2. A box (5.5 x 7.3 x 6.0 Å) was drawn surrounding L-Kyn to contain the calculation. The size of the box included the active site and extended beyond the narrowest part of the opening seen in Figure 3.17. Water molecules and substrate were deleted from the model prior to calculations. Calculations used a minimum cavity volume of 5 $Å^3$, a step size of 0.6 Å, an inner probe radius of 1.4 Å and an outer probe radius of 5.0 Å. The total surface area and total void volume in the box are reported.

For all other molecules the same size and position of box was necessary for comparison. This was achieved by first aligning the molecule in question to the *closed* molecule of substrate-bound dm2, as discussed earlier in this section. The box was drawn using the L-Kyn from the *closed* molecule of substrate-bound dm2. As before water molecules and substrate were deleted from the model prior to calculations.

Chapter 3: Structures of *Pf*KMO dm2

3.1. Introduction

As previously discussed in section 1.6.2 the original crystals of *Pf*KMO dm2 were grown using PEG 4K as the precipitant, and the conditions are detailed in Table 3.1.

Solution	Concentration
HEPES pH 7.0	0.1 M
PEG 4K	15 % w/v
Glycerol	10 % v/v
Propan-2-ol	10 % v/v
Dithiothreitol	2 mM
L-kynurenine	1 mM

Table 3.1 Conditions used to grow original crystals of dm2 and mutants. L-Kyn added only in the case where substrate-bound protein structures were desired.

These conditions led to either $P2_122_1$ crystals yielding a substrate-free structure with a maximum resolution of 2.26 Å or $I4_122$ crystals yielding a substrate-bound structure with a maximum resolution of 3.24 Å. The $I4_122$ crystals, in particular, had very high solvent content and very high B-factors overall. This clearly affected the quality and reliability of the resulting model.

In these structures of dm2 there was a C-terminus movement observed between the substrate-bound and the substrate-free forms, and it was hypothesised that the presence of L-Kyn caused this. The two forms, however, were in different space groups. It is therefore possible that the C-terminus movement was being caused by changes in the crystal contacts. If the C-terminus was being forced closed by the presence of other molecules one would expect to see a high number of contacts in the *closed* conformer and a lower number in the *open* conformer.

When mutant forms of dm2 were crystallised they showed a decreased movement in the C-terminus. The mutants still crystallised in the same space groups with $P2_122_1$ crystals yielding substrate-free structures and $I4_122$ crystals yielding substrate-bound structures. This occurrence seemed to refute the theory that crystal contacts caused the C-terminus shift.

It could be argued, however, that the presence of L-Kyn caused the protein crystals to grow in the different space groups, thus creating different crystal contacts and inducing the C-terminus movement.

Further analysis of the structures solved by Wilkinson was performed using *NCONT*[143, 169]. The program reported on the number of crystal contacts between the molecule in question, other protein molecules

in the ASU and symmetry generated molecules. Only residues after and including residue 370 were analysed for contacts closer than 5.0 Å. All internal contacts were discarded. The results are presented in Table 3.2.

Molecule	Total number of contacts
Substrate-bound dm2	
Molecule A	287
Substrate-free dm2	
Molecule A	332
Molecule B	530
Molecule C	491
Molecule D	251

Table 3.2 Total number of crystal contacts exhibited in dm2 structures grown in PEG 4K conditions.

The number of crystal contacts varies greatly from molecule to molecule even when in the same ASU. The substrate-free molecules show a variation of more than double the contacts solely from positioning within the ASU and the unit cell as a whole. Despite this the molecules show almost identical positioning of the C-terminus when overlaid with one another, as shown in Figure 3.1. The number of contacts in the substrate-bound structure is also fewer than most of the substrate-free molecules. These findings point towards crystal contacts having little influence upon the positioning of the C-terminus.

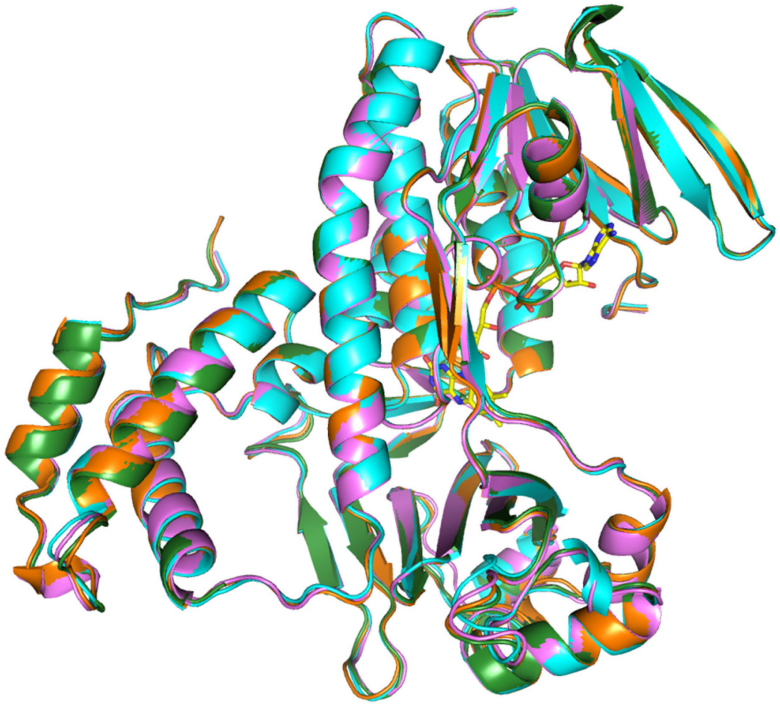

Figure 3.1 Crystal structure of substrate free dm2 grown in PEG 4K conditions by Wilkinson[143]. Molecules from the ASU are overlaid. RMSD values were generated by aligning each molecule to molecule A (cyan). RMSD values of the overlays range from 0.458-0.730 Å, from at least 438 aligned Cα atoms in each case.

Using new crystallisation conditions this chapter reports high-resolution structures of dm2 in both substrate-bound and substrate-free forms. Interactions between binding residues and L-Kyn are examined, as is the movement and putative bonding network between the C-terminus and the rest of the protein.

3.2. Substrate bound dm2 crystal structure

3.2.1. Protein purification

The dm2 protein was expressed and purified as described in Chapter 2. An example of a typical chromatogram from both the Q Sepharose column and the Superdex 75 column are shown in Figure 3.2 and Figure 3.3, respectively.

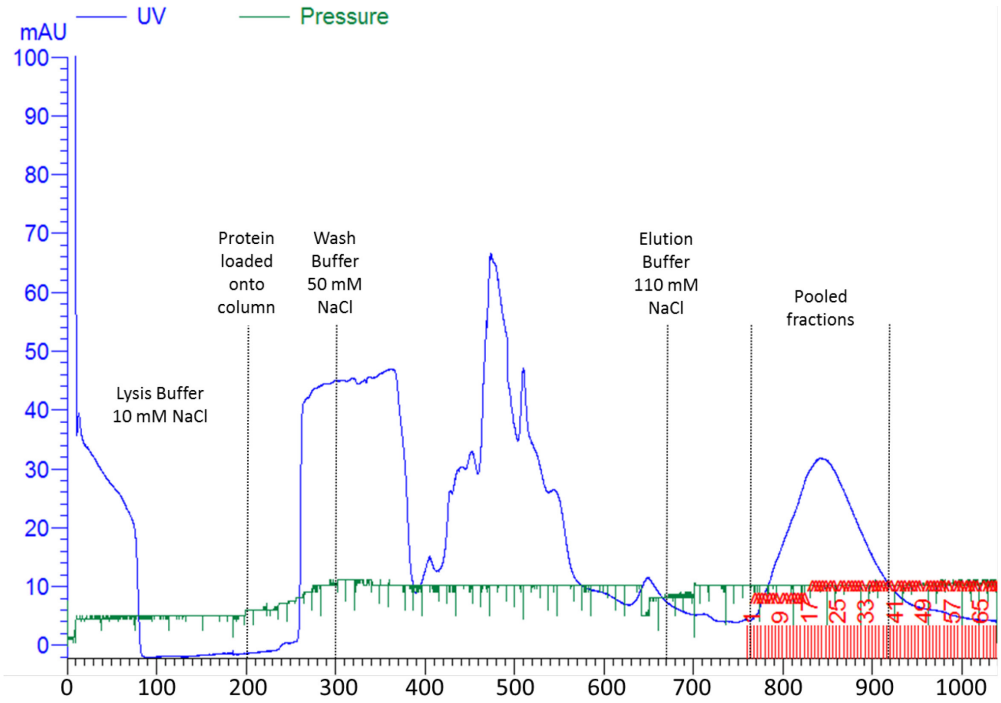

Figure 3.2 Typical chromatogram of absorbance at 280 nm of eluent from the Q Sepharose column against volume in ml. The column was pre-equilibrated with lysis buffer; the chromatogram shows when the protein sample and buffers were added to the column. For further details refer to section 2.4.1.

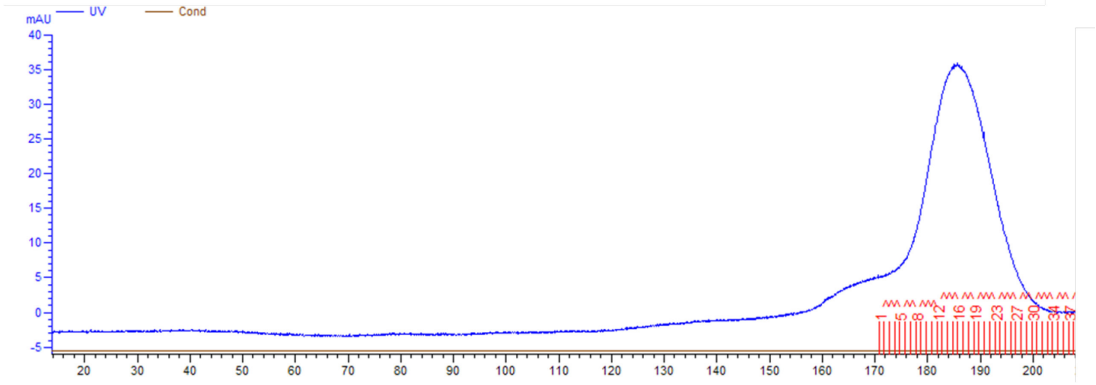

Figure 3.3 Typical chromatogram of absorbance at 280 nm of eluent from the Superdex 75 column against volume in ml. The protein was collected by an automated fraction collector monitored by UV absorbance at 280 nm. For further details refer to section 2.4.1.

The fractions were tested for purity and concentration using absorbance measurements described in section 2.5.2 and pooled. SDS-PAGE was performed of samples from throughout the purification process. A typical SDS-PAGE gel is shown in Figure 3.4.

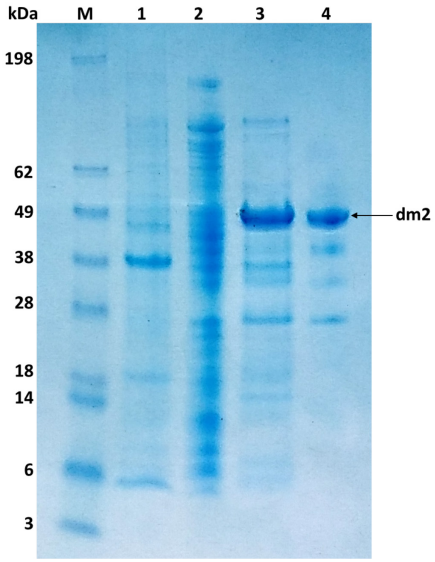

Figure 3.4 Typical SDS-PAGE gel of dm2 purification. novex® SeeBlue® Pre-Stained Standard was used as a marker (M); the weights of the standards are included in kiloDaltons (kDa). Lane 1 – Cell pellet of centrifuged lysed cells. Lane 2 – Supernatant of centrifuged lysed cells. Lane 3 – Pooled fractions of dm2 after the Q Sepharose column. Lane 4 – Pooled fractions of dm2 after the Superdex 75 column.

The protein was concentrated by centrifugation to a concentration of 675 μM, confirmed by UV-Vis spectrophotometry, shown in Figure 3.5.

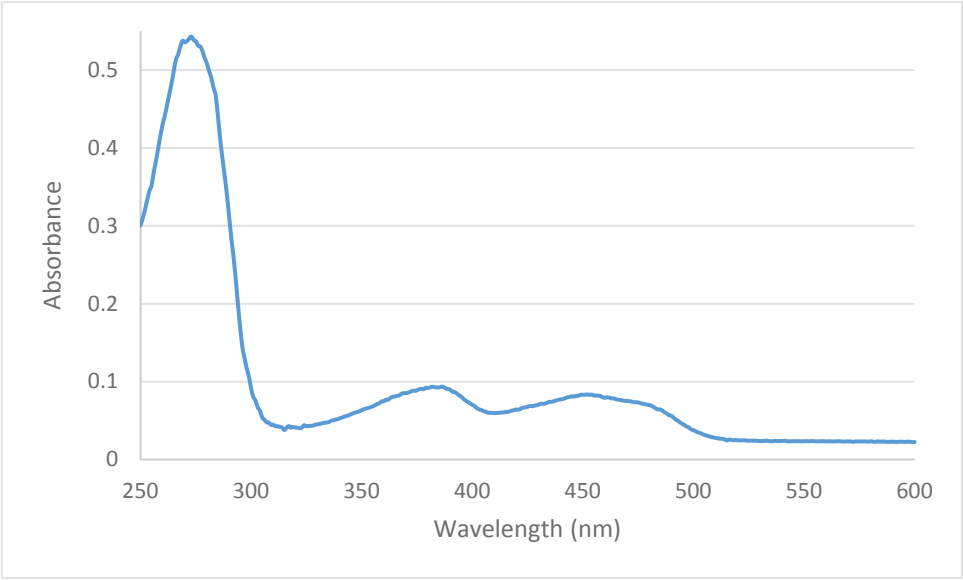

Figure 3.5 UV-visible spectrum of 1:100 dilution of dm2. Concentration calculated from absorbance of 0.083 at 450 nm with molar extinction coefficient (ε_{450}) of 12,300 M^{-1} cm^{-1}.

The Michaelis-Menten kinetics of the dm2 was assessed using the assay outlined in section 2.6. The results were analysed with the program *Origin 2015* (OriginLab, Northampton, MA).

An example of a Michaelis-Menten plot is shown in Figure 3.6.

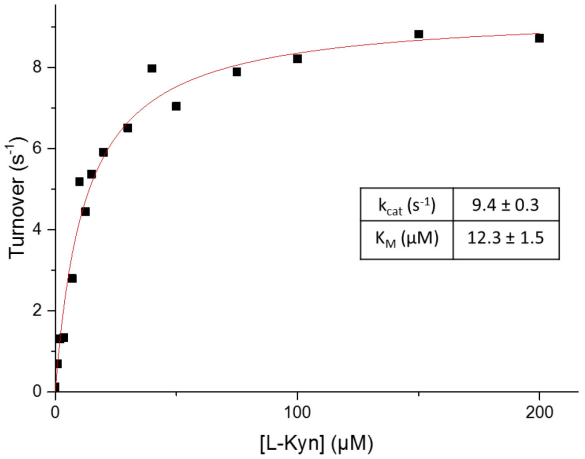

Figure 3.6 Graph of a typical Michaelis-Menten plot for dm2 using L-Kyn as substrate, showing K_M and k_{cat} values.

The resulting protein was deemed pure enough and presented an activity profile to allow crystallisation.

3.2.2. Crystal growth and data collection

New crystallisation conditions arose as part of an industrial collaboration[173]. The conditions were initially discovered from sparse matrix screen tests using the JCSG Core Suite III and the JCSG+ screens developed by the Joint Centre for Structural Genomics[174, 175]. These will be referred to as the PEG 8K conditions and are described in Table 3.3.

Solution	Concentration
Na cacodylate pH 6.5	0.08 M
PEG 8000	10-20 % w/v
Ca acetate	0.16 M
Glycerol	20 % v/v
L-kynurenine	1 mM

Table 3.3 Conditions used to grow improved crystals of dm2. L-Kyn added only for substrate-bound crystals. w/v is weight per volume, v/v is volume per volume.

The new PEG 8K conditions were used to set up hanging drop trays as described in Chapter 2. Two trays were set up, one co-crystallised with 1 mM L-Kyn, the other without L-Kyn. Within 24 hours a number of large crystals had formed as seen in Figure 3.7. Selected crystals were then soaked in 1 μL drops comprising mother liquor stabilised with 25 % PEG8K and 1 mM L-Kyn. Crystals were then left for a further 24 hours.

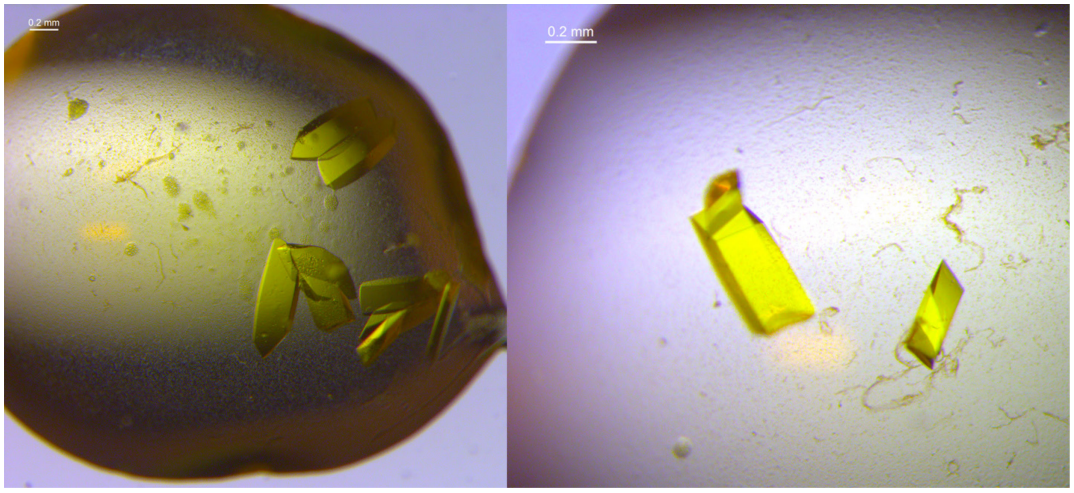

Figure 3.7 Crystals of dm2 grown in PEG 8K conditions. The camera provided a 0.2 mm scale bar.

Individual crystals were flash-cooled in liquid N_2 as described in section 2.7.4 and sent to the Diamond Light Source. Using beamline i04-1 data were collected. Using Fast single data reduction pipeline software (FASTDP) the space group of all the crystals analysed was determined as $P12_11$. The unit cells of the crystals, both substrate bound and substrate free, were all very similar; the approximate values are shown in Table 3.4.

Crystal Parameter	**a**	**b**	**c**	**β**
Approximate value	70 Å	52 Å	138 Å	104°

Table 3.4 Values of unit cell parameters for crystals grown in PEG8K conditions.

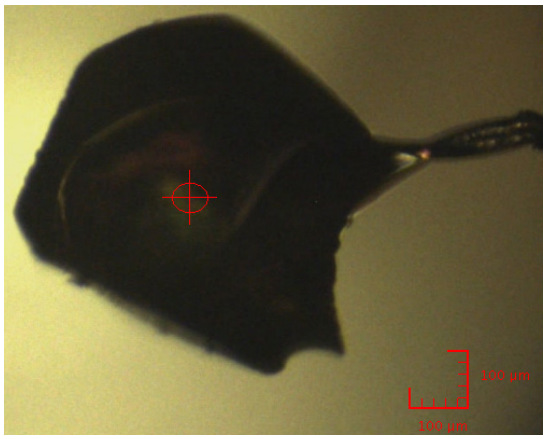

Figure 3.8 Mounted crystal of substrate-exposed dm2 used for data collection.

From the crystal shown in Figure 3.8 1800 images were collected with beamline transmission 18.75 %, exposure of 0.05 s and oscillation of 0.10°. One of these images is shown in Figure 3.9.

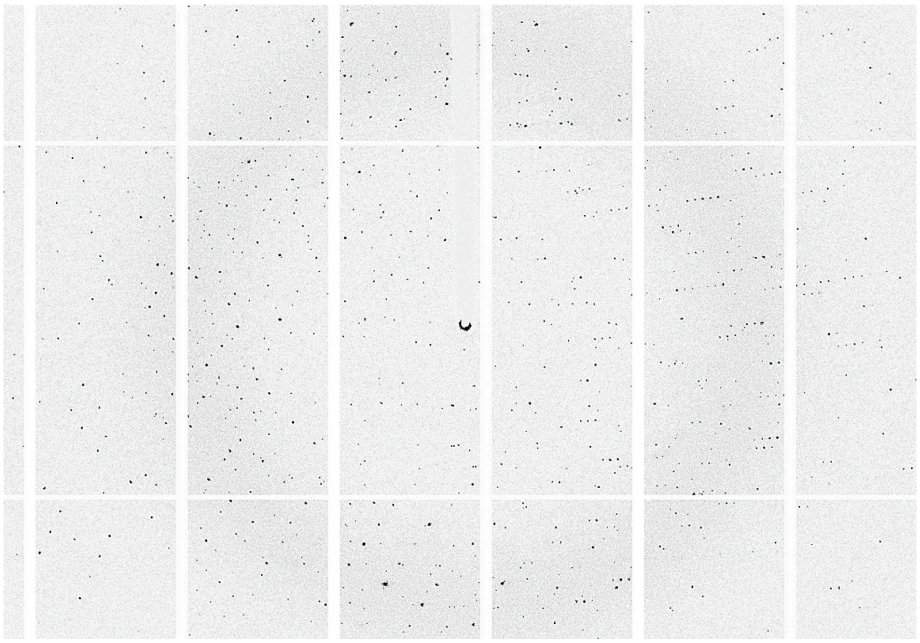

Figure 3.9 Diffraction image from dm2 crystal exposed to L-Kyn.

The statistics shown in Table 3.5 were calculated using the programs *iMosflm* and *Scala*[159, 160].

Resolution (Å)	R_{merge}	Mean (I)/σ(I)	Completeness (%)	Multiplicity
67.8-1.5	0.062	8.1	99.7	3.1

Table 3.5 Data collection statistics for a substrate-exposed dm2 crystal. All calculated from overall data, for full data collection statistics see Appendix 7.1.1.

Molecular replacement was used to calculate structure factors. For further details see section 2.8. The structure was solved with two dm2 molecules in the ASU and a solvent content of 49.5%. As detailed in section 2.8, *Refmac5* was used to carry out restrained refinement and *Coot* was used to edit the model[163, 164]. An overall image of the two molecules in the ASU can be seen in Figure 3.10.

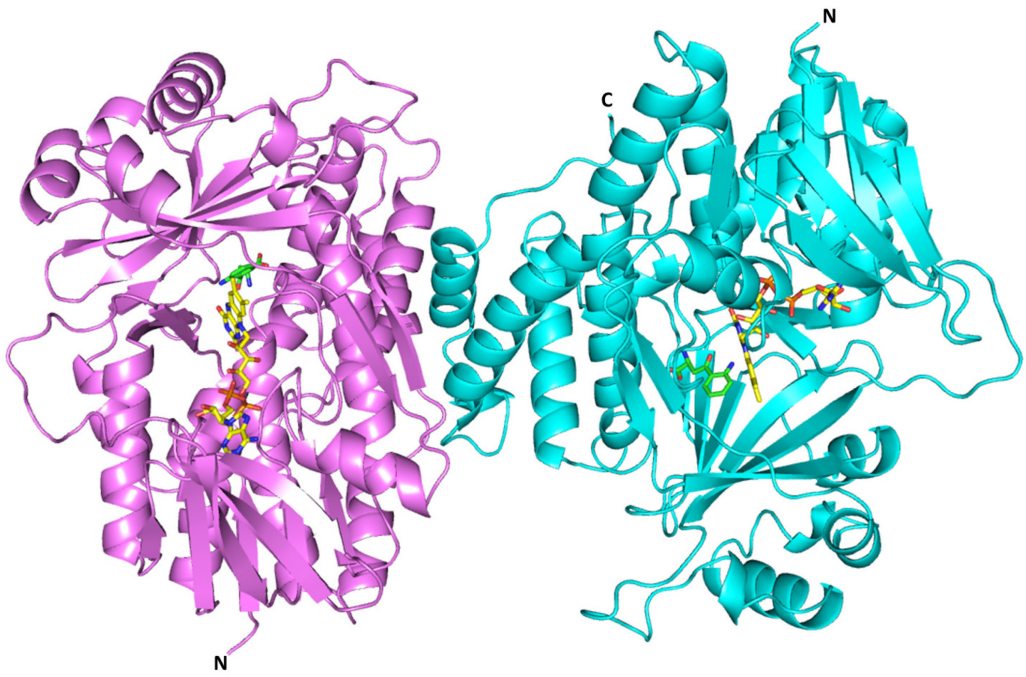

Figure 3.10 Crystal structure of substrate-bound dm2 showing both molecules in the ASU. Molecule A is coloured cyan, molecule B is coloured violet.

3.2.3. FAD and Kynurenine Binding

The first refinement of the model showed clear electron density for the FAD cofactor in both molecules, as shown in Figure 3.11.

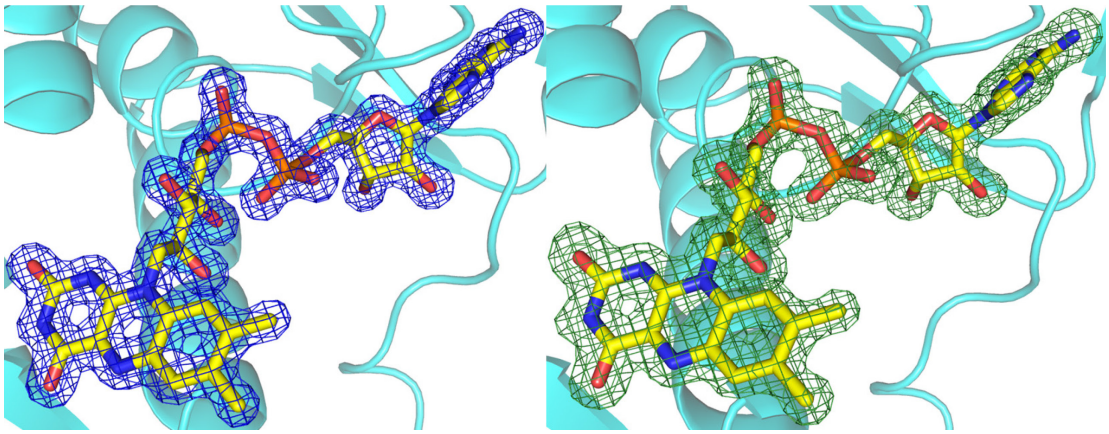

Figure 3.11 FAD bound in dm2 with the 2Fo-Fc map on the left in blue (contoured at 0.51 eÅ^{-3}, 1.51 σ) and with the Fo-Fc map, calculated in the absence of FAD, on the right in green (contoured at 0.35 eÅ^{-3}, 3.0 σ).

The model contained residues 7-457 in molecule A and residues 7-456 in molecule B. *Refmac5* was used towards the end of refinement to locate 1074 water molecules within the model. These were then manually removed from the active site of both molecules to investigate the presence of ligand electron density. When refined once more there was clear electron density matching the shape of L-Kyn in

molecule A. The other molecule showed less definitive evidence of L-Kyn binding, and this could only be visualised at low levels of contouring. Both active sites can be seen in Figure 3.12.

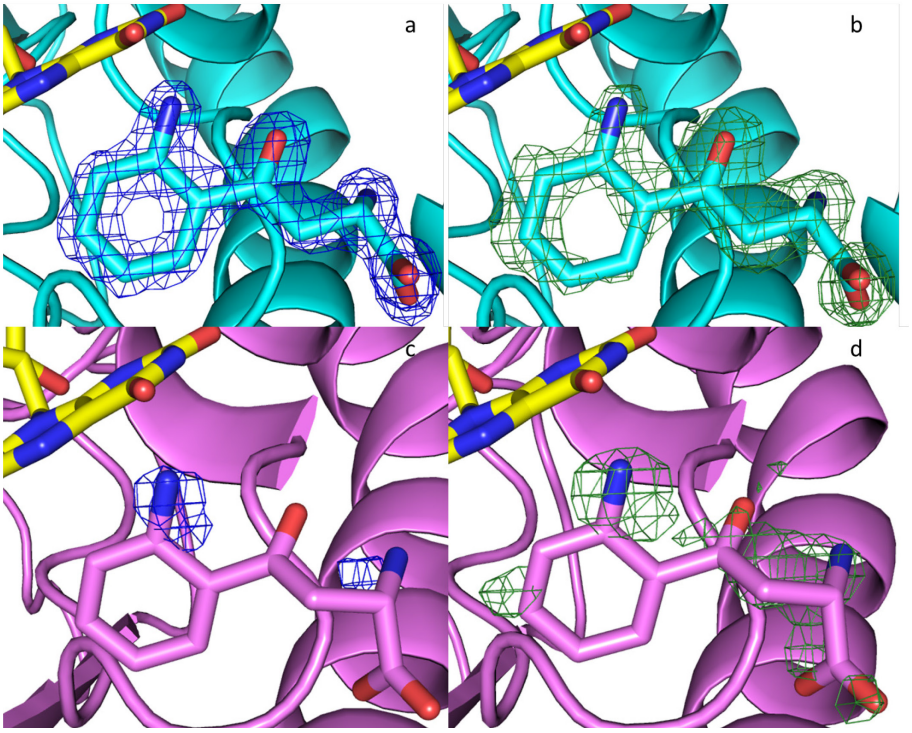

Figure 3.12 L-Kyn modelled in the binding site of dm2 **a.** 2Fo-Fc map for molecule A **b.** Fo-Fc map for molecule A **c.** 2Fo-Fc map for molecule B **d.** Fo-Fc map for molecule B. 2Fo-Fc maps are shown in blue (contoured at 0.51 eÅ^{-3}, 1.51 σ). Fo-Fc maps, calculated in absence of L-Kyn, are shown in green (contoured at 0.35 eÅ^{-3}, 3.0 σ).

The final model was refined to give R and R_{free} values of 0.2328 and 0.2715, respectively, to a resolution of 1.5 Å. The full statistics can be seen in Appendix 7.1.1.

3.2.4. Comparison of new and old dm2 structures

The previously solved substrate-bound dm2 model, grown in PEG 4K conditions, was compared to molecule A of the new model. This was done using RMSD as detailed in section 2.8. Both are very similar overall (RMSD = 1.492 Å from 450 aligned Cα atoms), with FAD placement almost identical and the secondary structures overlapping very closely, as shown in Figure 3.13.

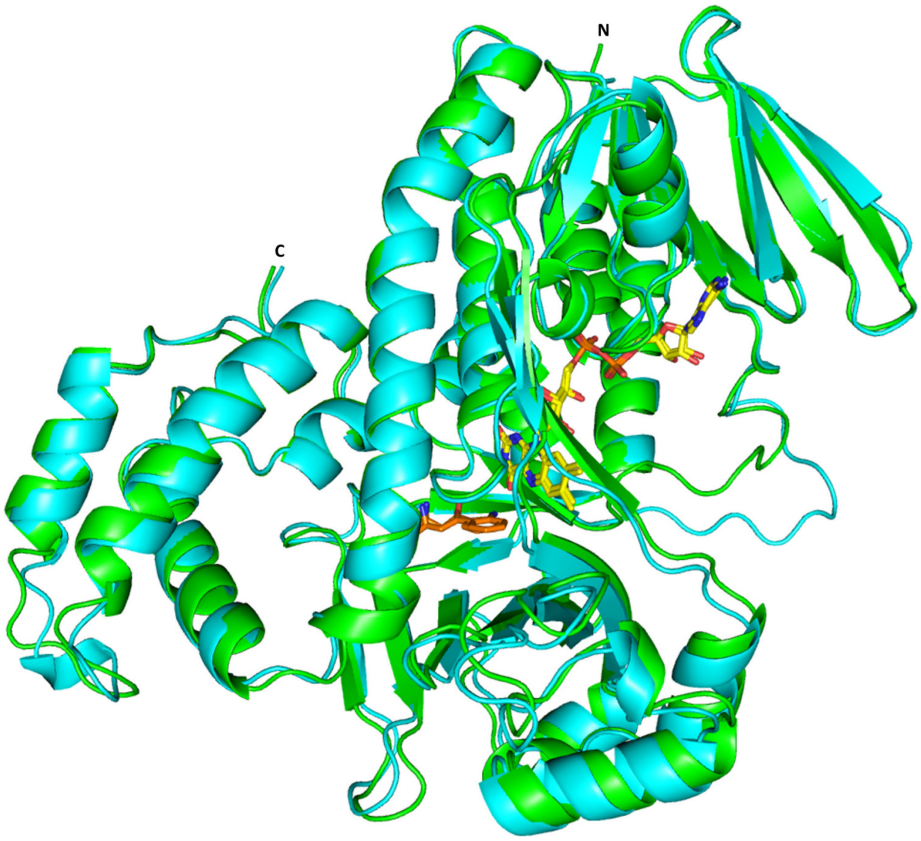

Figure 3.13 Overlaid crystal structures of substrate bound dm2 grown in original and improved conditions. Original molecule shown in green, molecule A of improved PEG 8K conditions shown in cyan. FAD is shown in yellow, L-Kyn is shown in orange. RMSD = 1.492 Å from 450 aligned Cα atoms.

The positioning of the L-Kyn within the active site of the new structure is shown in Figure 3.14. There are small differences in the protein-ligand interactions between this structure and the original dm2 model. These can be seen in Figure 3.14.

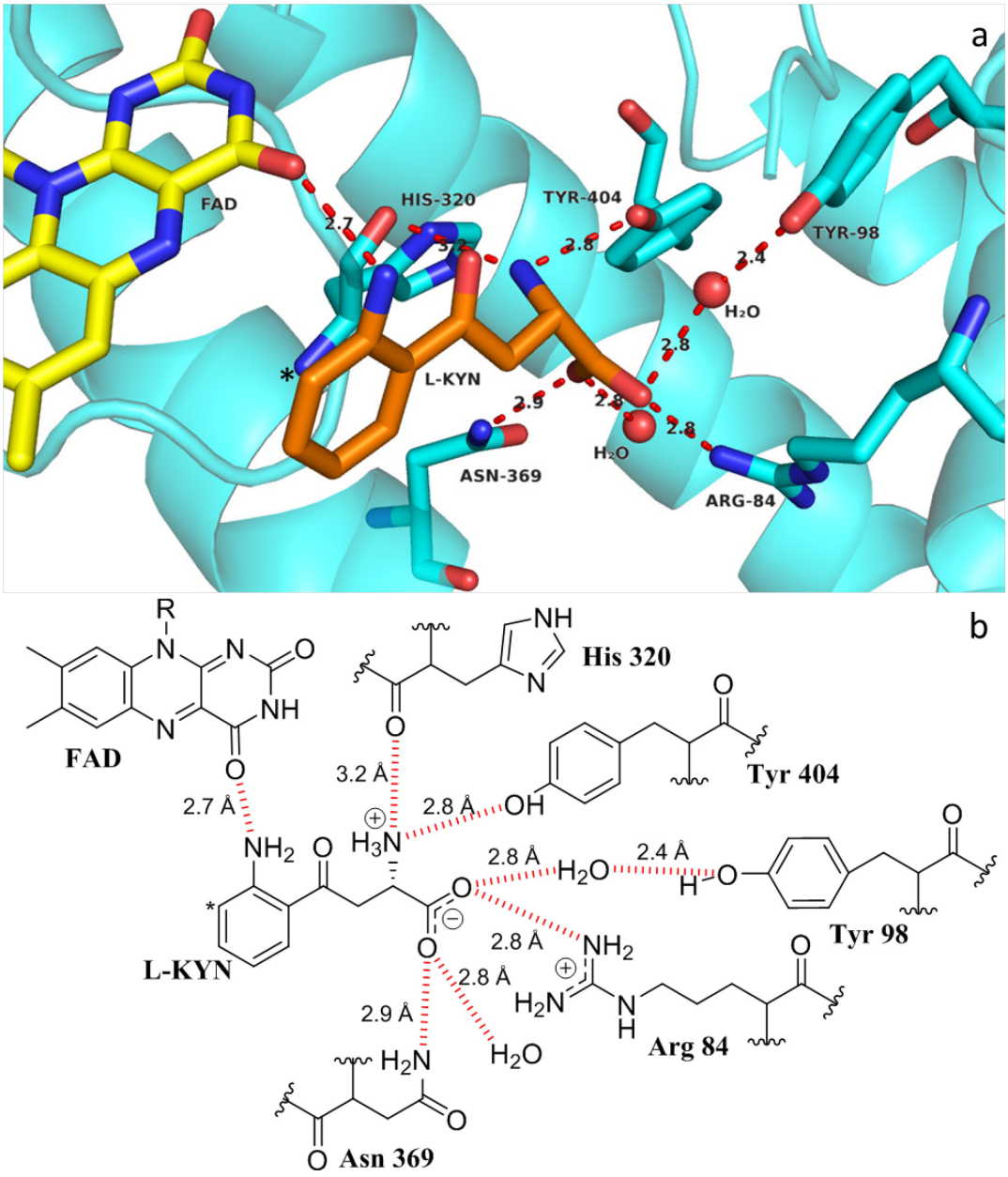

Figure 3.14 a. Crystal structure of active site of dm2 bound with L-Kyn, molecule A. **b.** Chemdraw schematic of dm2 bound with L-Kyn. Distances between atoms are in Å and shown as red dashed lines. The site of oxidation is denoted with an asterisk (*).

In this model L-Kyn is shown to form H-bonds with many of the same residues as previously seen. For example the FAD carbonyl to L-Kyn's aromatic amine and the N369 to the carboxylic acid interactions are the same, with similar distances between the groups. R84 shows similar interactions to those observed in previous structures. The positive guanidinium group of R84 forms an electrostatic interaction with the negative carboxylate group of L-Kyn, implying a stronger interaction than with the other binding residues. L-Kyn's aliphatic amine forms H-bonds between H320 and Y404, neither of which were seen in previous models. In contrast, in this structure Y98 is too far from the substrate to

show a direct interaction. Instead both the hydroxyl of Y98 and the carboxylate of L-Kyn form an H-bond to the same water molecule. There is a water molecule seen in the same relative position in both molecules of the ASU. A water molecule is also seen interacting with the carboxylate; an equivalent water molecule is present in the other protein molecule. The hydrophobic residues within the active site, such as I224 and F238, form a lipophilic area surrounding the aromatic ring of L-Kyn (not pictured). The overall position of L-Kyn allows for it to react with the flavin-C4a-hydroperoxide group once it has formed.

This model reinforces R84 as a key-binding residue and further highlights the importance of Y404. The role of Y98 is proposed to be less influential than assumed from previous models.

3.2.5. Comparing substrate bound dm2 molecules within the asymmetric unit

Comparing the active site within the two molecules in Figure 3.15 there is very little difference in positioning of residues. There is a small difference in R84 positioning between the two molecules. This may be influenced by the electrostatic interaction with the substrate, pulling the side chain into position.

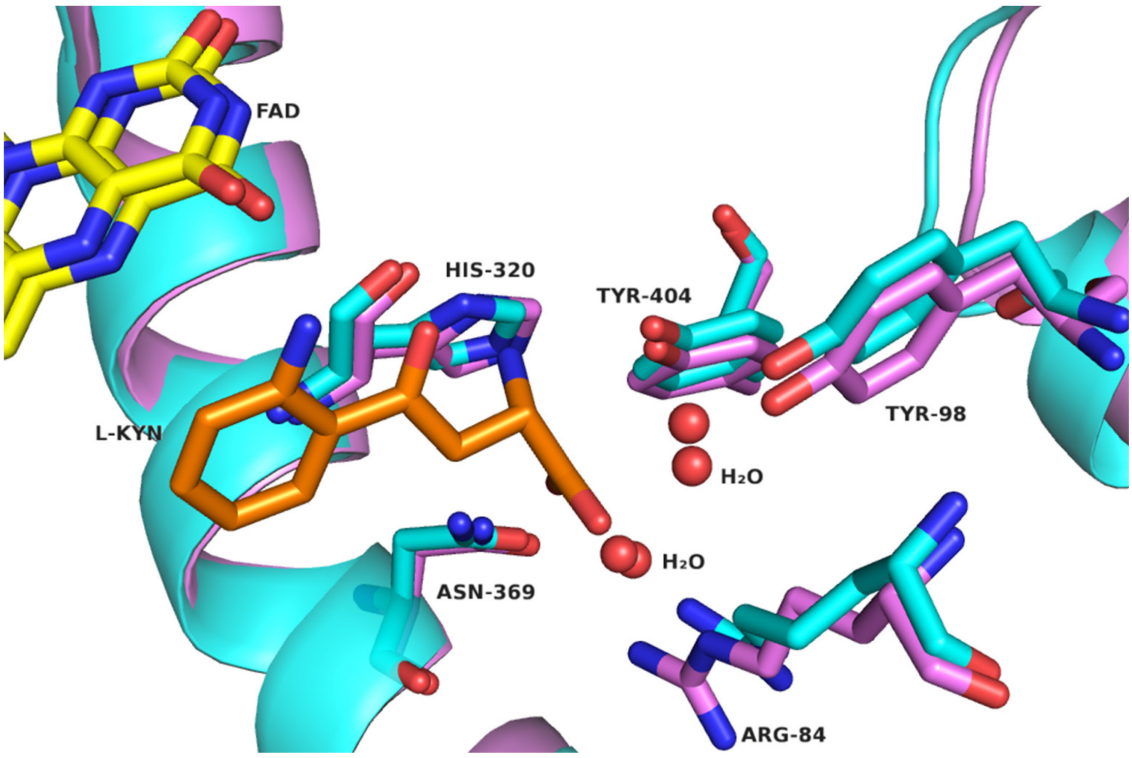

Figure 3.15 Crystal structure of dm2 with L-Kyn bound, showing overlaid active sites of molecule A (cyan) and molecule B (violet). RMSD = 1.165 Å from 436 aligned Cα atoms.

The active sites of both molecules within the ASU of the PEG 8K crystals are the same within experimental error. There are, however, large differences between the two molecules overall. In the $I4_122$ form only one molecule was present in the ASU, with 100% occupancy of substrate. In the $P12_11$ form the occupancy of substrate differs significantly between the molecules in the ASU, as seen in

Figure 3.12. The molecules themselves also differ. When overlaid with one another, molecules A and B of the $P12_11$ dm2 show a familiar movement of the C-terminal domain, despite a low RMSD. This is shown in Figure 3.16.

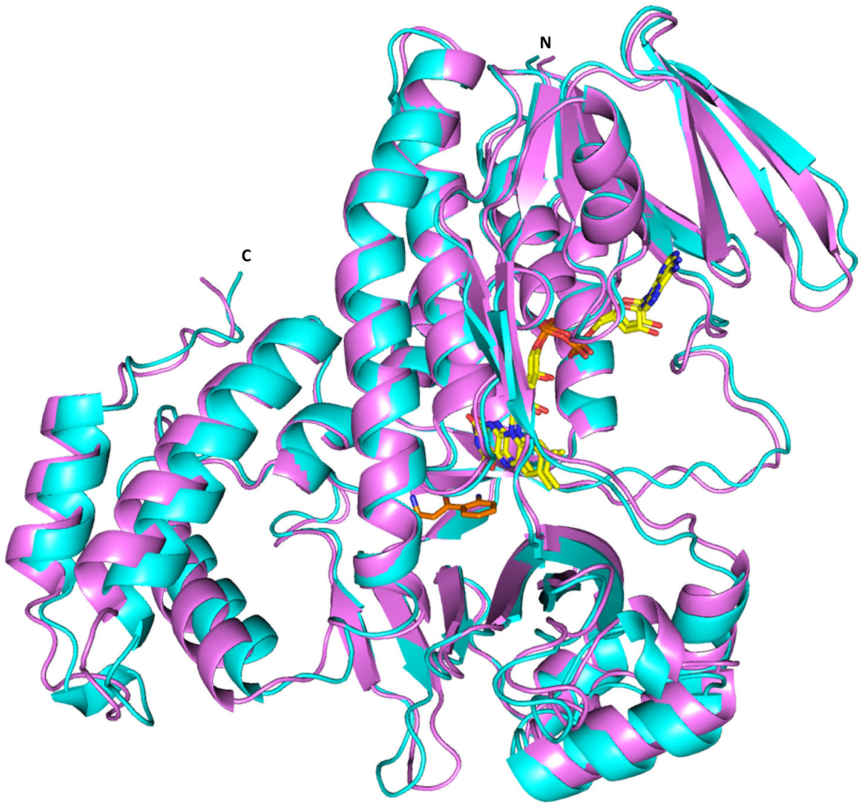

Figure 3.16 Crystal structure of dm2 bound with L-Kyn, molecules A and B overlaid with one another. Molecule A is shown in cyan, molecule B is shown in violet. RMSD = 1.165 Å from 436 aligned Cα atoms.

The C-terminal movement matches the movement discussed in section 1.6.3 purportedly caused by the presence of L-Kyn. Molecule A matches with the *closed* conformation, while molecule B matches the *open* conformation. Interestingly the occupancy level of substrate matches with these positions, the *closed* molecule A shows high occupancy of substrate, while the *open* molecule B shows very low occupancy of substrate. The channel leading to the active site is also seen in the *open* position, as shown in Figure 3.17. This would allow L-Kyn to diffuse in and out of the enzyme, explaining its low occupancy level. The access route is shut in *closed* molecule A, causing L-Kyn to remain trapped in the active site and increasing its occupancy in the model.

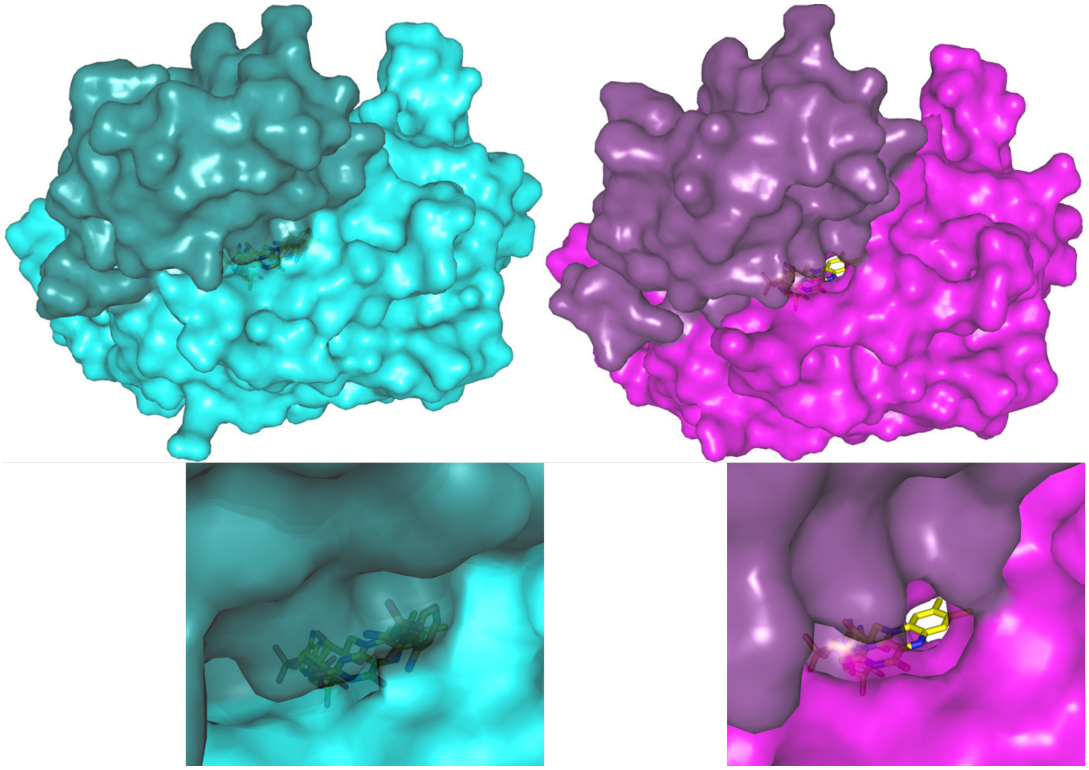

Figure 3.17 Crystal structure of dm2 bound with L-Kyn showing surfaces. Molecule A is shown in cyan on the left, molecule B is shown in violet on the right. The C-terminus is shown in darker shade. An enlarged view of the ligands is also shown.

Surface calculations were performed on the two molecules using the program *KVFinder*[172]. The results are shown in Table 3.6.

	Void Volume (Å^3)	Surface Area (Å^2)
Closed molecule	376	413
Open molecule	374	431

Table 3.6 Values of surface calculations of substrate-bound dm2.

The volume of the active site seems very similar for both molecules. The surface area, however, is shown to be slightly larger in the *open* molecule. This is to be expected since there is a larger opening seen in the *open* molecule.

A slight movement between the two molecules is also observed in the bottom right of Figure 3.16, corresponding to residues 240-290. The B-factors of these residues are consistently higher than the mean indicating they are more mobile than the overall structure. It is possible that their movement is linked to the movement of the C-terminus. Given this is the first observation of this movement its relevance to the *open* and *closed* conformations cannot be commented on until we examine more structures. This movement is not seen in the previous models grown in PEG 4K conditions. The residues

in question do not seem involved in the active site, nor the opening of the channel shown in Figure 3.17. They have also never been recognised in previous literature as important.

3.3. Substrate free dm2 crystal structure

3.3.1. Crystal growth and data collection

Having been expressed and purified as detailed in section 3.2.2, dm2 was crystallised using the PEG 8K conditions but at no point was the enzyme exposed to L-Kyn. Individual crystals were flash-frozen in liquid N_2 and sent to the Diamond Light Source. Using beamline i02 data were collected. Using FASTDP the space group of all the crystals analysed was determined as $P12_11$ with the same approximate unit cell dimensions as detailed in Table 3.4. From the crystal shown in Figure 3.18 2000 images were collected with beamline transmission 20.00 %, exposure of 0.05 s, and oscillation of 0.10°.

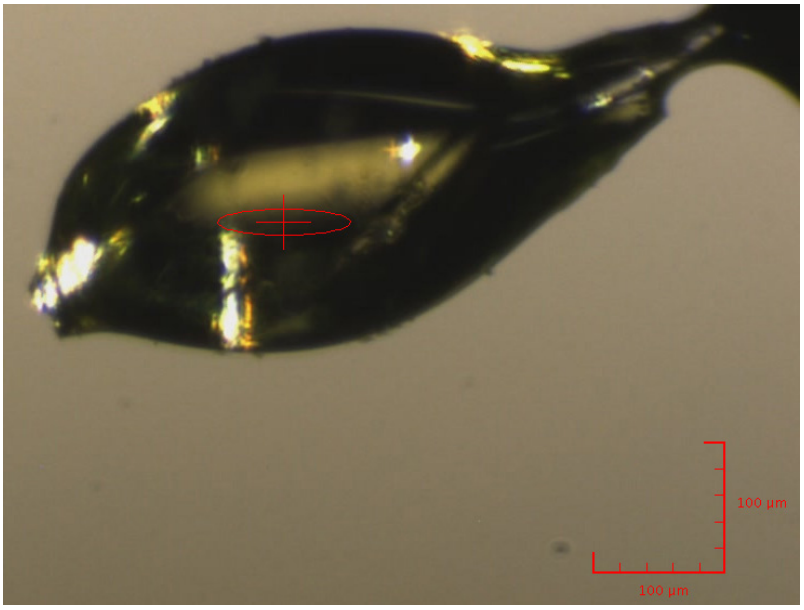

Figure 3.18 Mounted crystal of substrate free dm2 used for data collection.

The statistics shown in Table 3.7 were calculated using *iMosflm* and *Scala*[159, 160].

Resolution (Å)	R_{merge}	Mean (I)/σ(I)	Completeness (%)	Multiplicity
67.76-1.79	0.047	14.3	99.6	3.4

Table 3.7 Data collection statistics for a substrate-bound dm2 crystal. All calculated from overall data, for full data collection statistics see Appendix 7.1.1.

Molecular replacement was used to calculate structure factors. For further details see section 2.8. The structure was solved with 2 molecules in the ASU and a solvent content of 48.06%. As detailed in

section 2.8, *Refmac5* was used to carry out restrained refinement and *Coot* was used to edit the model[163, 164]. Both molecules within the model contain residues 6-461. 532 waters were located using *Refmac5* and the R and R_{free} values of the final model were 0.1692 and 0.2080, respectively, to a resolution of 1.79 Å. The full statistics can be found in Appendix 7.1.1. An overall image of the model is shown in Figure 3.19.

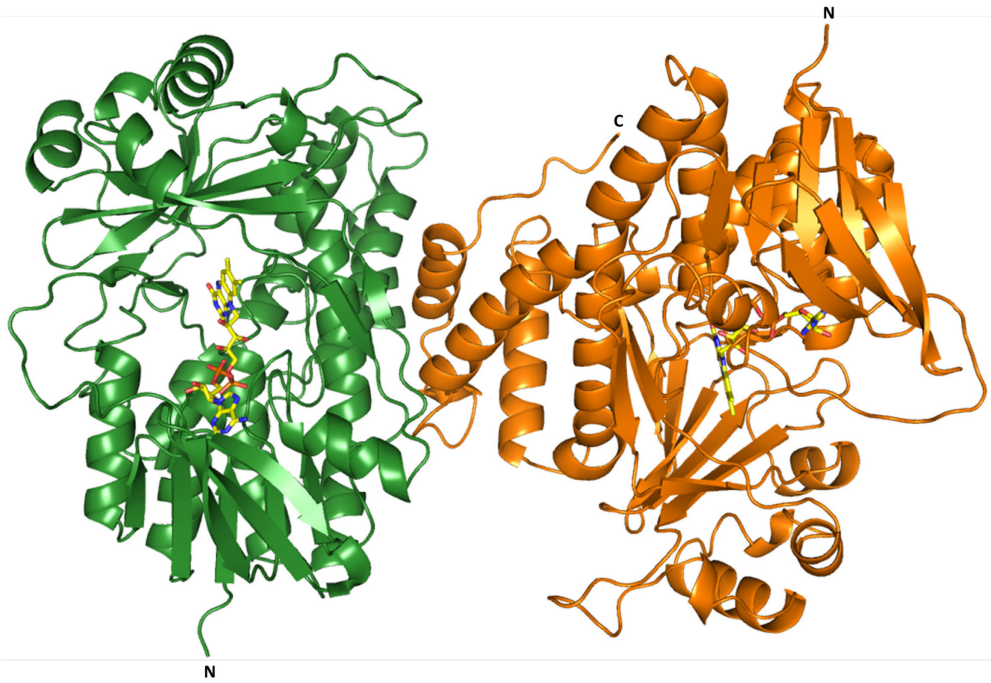

Figure 3.19 Crystal structure of substrate-free dm2 showing both molecules in the ASU. Molecule A is coloured orange; molecule B is coloured green.

3.3.2. Overall structure of substrate free dm2

The most immediate finding is that the two molecules also exhibit *open* and *closed* positions. There is a very similar shift in the C-terminal domain to that observed in the substrate-bound model, as shown in Figure 3.20.

Figure 3.20 Crystal structure of L-Kyn free dm2, molecules A and B overlaid with each other. RMSD = 1.028 Å from 444 aligned Cα atoms.

The channel opening and closing is also observed, as shown in Figure 3.21. The *closed* position of molecule A shows a very slight opening to the active site whereas the *open* molecule B shows a channel similar in size to those seen in previous models.

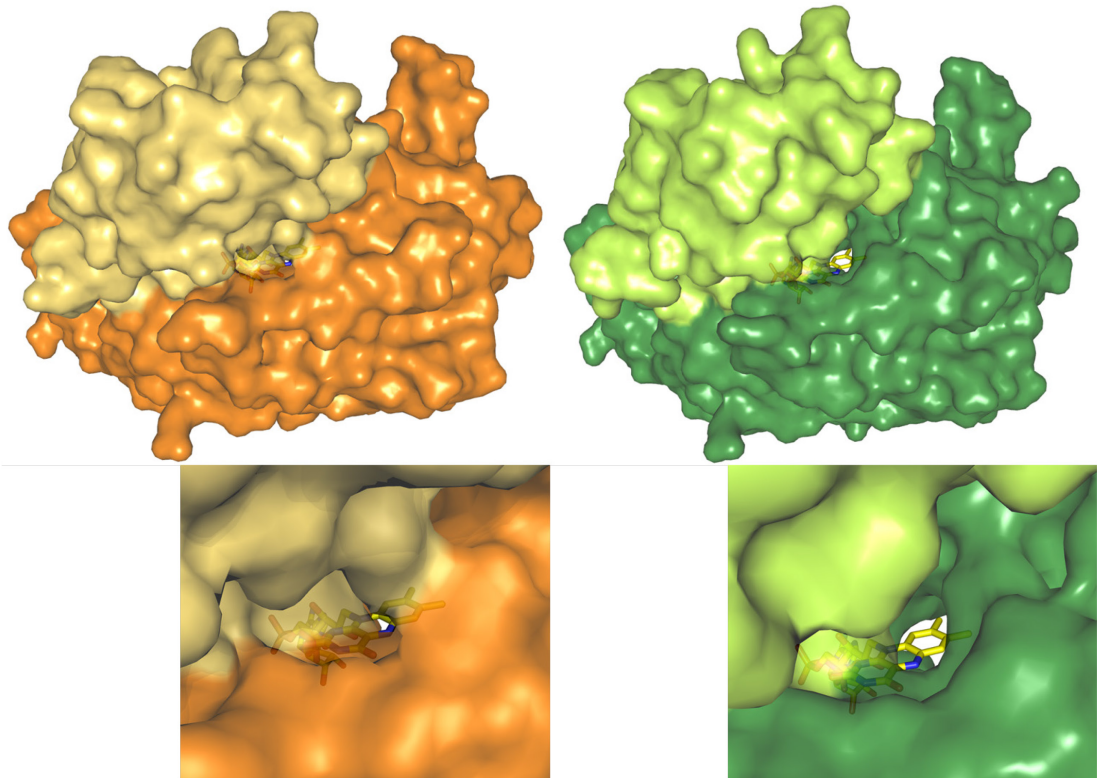

Figure 3.21 Crystal structure of L-Kyn free dm2, showing surfaces. Molecule A is shown in orange on the left, molecule B is shown in green on the right. The C-terminus is shown in lighter shade. An enlarged view of the ligands is also shown.

As with the substrate-bound model, surface calculations were performed on the active site of each molecule. The results are shown in Table 3.8.

	Void Volume (Å^3)	Surface Area (Å^2)
Closed molecule	383	448
Open molecule	322	382

Table 3.8 Values of surface calculations of substrate-free dm2.

These results are contradictory to our expectations of the more *open* molecule having a larger surface area or void volume in the active site. These findings, as well as the appearance of a small opening in the *closed* molecule, may indicate a more fluid transience between conformations.

In the PEG 8K conditions the similarity, in terms of crystal symmetry and unit cell dimensions, between the substrate bound and substrate free forms of dm2 allows for a better comparison than previously possible. An overlay of the two overall models is shown in Figure 3.22.

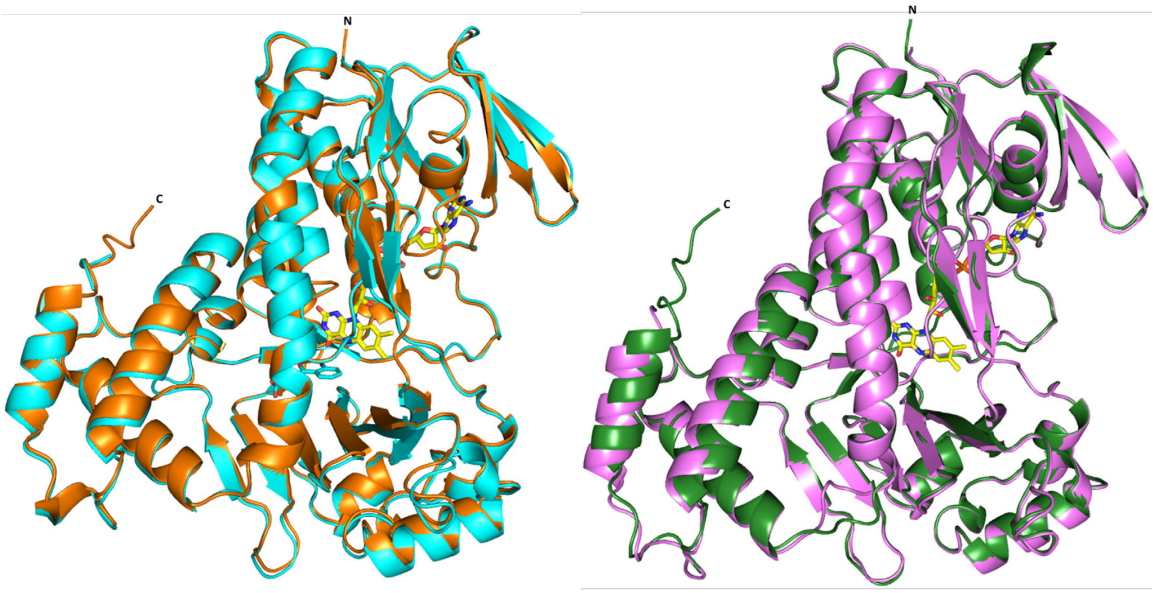

Figure 3.22 Crystal structures of dm2. L-Kyn bound dm2 is shown in cyan and violet, L-Kyn free dm2 is shown is orange and green. Molecules A are shown on the left with RMSD = 0.514 Å from 438 aligned Cα atoms. Molecules B are shown on the right with RMSD = 0.731 Å from 445 aligned Cα atoms.

The equivalent molecules of both crystals align very closely with one another, with small RMSD values. The C-termini for both pairs of molecules are in very similar positions. This shows that dm2 is in both the *open* and *closed* conformations regardless of substrate exposure. The fact that dm2 is able to take up the *closed* position in the absence of L-Kyn contradicts the hypothesis that the C-terminus shift is solely controlled by the presence of substrate.

The crystal contacts of the structures presented in the chapter were analysed with *NCONT* as discussed in section 2.8.1. The results are shown in Table 3.9.

Molecule	**Total number of contacts**
Substrate-bound dm2	
Open molecule	344
Closed molecule	317
Substrate-free dm2	
Open molecule	437
Closed molecule	443

Table 3.9 Total number of crystal contacts exhibited in dm2 structures grown in PEG 8K conditions.

Much like the crystal contact analysis in section 3.1; the number of contacts does not seem to affect the positioning of the C-terminus. In the substrate-bound structure the *open* molecule has much more

contacts, whereas the substrate-free structure shows slightly more contacts in the *closed* molecule. This continues the trend of the number of contacts bearing no relation to the positioning of the C-terminus.

The movement of residues 240-290 is also observed in Figure 3.15 and is similar to that seen in the substrate-bound structure. The residues align to each other in Figure 3.22, much like the C-terminus. This implies that the movement of these residues is connected to the movement in the C-terminus. The movement to the *closed* conformation may cause residues 240-290 to move to accommodate the change. If this were a necessary movement, however, it would be expected to occur in the previous models grown in PEG 4K conditions. The link, therefore, remains unclear but their putative connection will be scrutinised further throughout this work. The reasoning for observing specifically residues 240-290 is because they denote the most mobile residues of this section of the protein. Outwith these residues the movement of residues observed between molecules is minimal.

When comparing the active sites of the two substrate-free molecules we see very little difference; much like the substrate-bound model in section 3.2.5. This is shown in Figure 3.23.

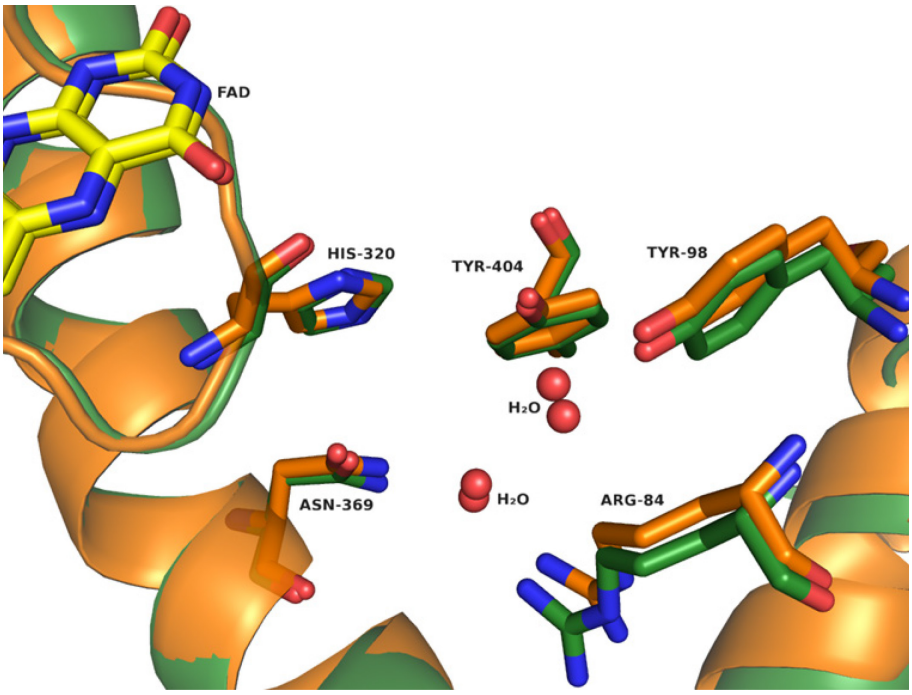

Figure 3.23 Crystal structure of substrate-free dm2, showing overlaid active sites of molecule A (orange) and molecule B (green). RMSD = 1.028 Å from 444 aligned Cα atoms.

Residues within the active site of equivalent molecules from both substrate-free and substrate-bound forms are also very similar, as shown in Figure 3.24.

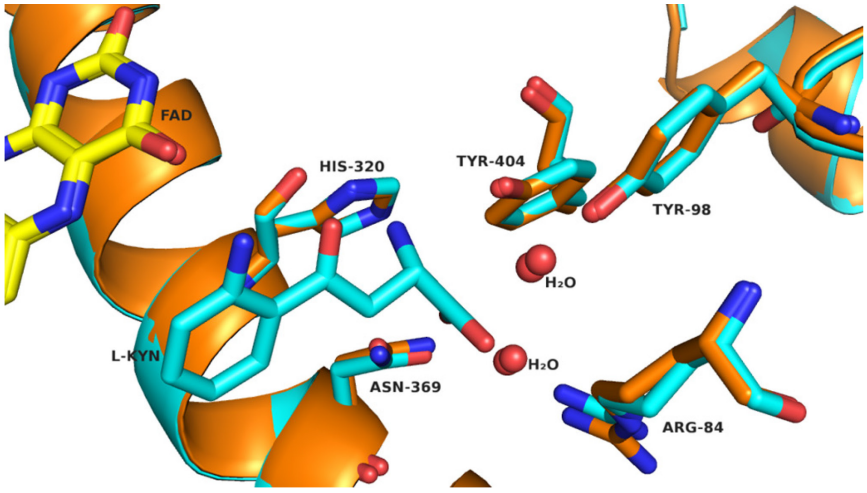

Figure 3.24 Crystal structure of overlaid molecules A from substrate-free and substrate-bound dm2, showing the active site. Substrate-bound dm2 is shown in cyan, substrate-free dm2 is shown in orange. RMSD = 0.514 Å from 438 aligned Cα atoms.

The similarity between the two molecules within each crystal allows us to observe that all four molecules, both substrate-bound and -free, have very similar positioning of binding residues. The lack of difference, especially in residue R84 given its potential role within the C-terminus movement, implies the absence of any induced fit causing the enzyme to change conformation. This adds to the theory that L-Kyn is not responsible for the shift in the C-terminus.

3.4. C-terminus movement

The above findings do not diminish the importance of the C-terminus in substrate binding. The highest occupancy of L-Kyn is observed in a molecule when it is in the *closed* conformation. This section discusses the positioning of the residues identified as having a role in the possible bonding network between the C-terminus and the rest of the protein.

The C-terminus of dm2 is relatively mobile as shown by its high B-factors. In order to compare B-factors the mean B-factors for all atoms in residues 370-400 of each molecule were divided by the mean B-factor for all atoms in whole structure. This is therefore an indication of relative mobility in residues 370-400 compared to the mobility of the overall structure. The residues 370-400 were chosen as they bookend the most mobile region of the C-terminal domain. Residue 370 is at the end of a relatively immobile α-helix and the start of a flexible loop. At residue 400 the B-factor value drops significantly as proximity to the active site increases, indicating the end of the mobile region. The values of this calculation for each molecule in both substrate-bound and substrate-free forms of dm2 are shown in Table 3.10.

Molecule	Mean B-factor of C-terminus/Mean B-factor of structure
Bound closed (A)	1.2
Bound open (B)	1.9
Free closed (A)	1.3
Free open (B)	1.8

Table 3.10 Average B-factor for the residues 370-400 of substrate-free and substrate-bound forms of dm2 divided by average B-factor of overall structure.

In both of the models the *open* conformation has a higher average B-factor in the C-terminus. This implies that the *open* form is more mobile and disordered when compared to the *closed* form. When comparing the molecules within each ASU the substrate-free molecules are slightly closer in mobility than their substrate-bound counterparts. This could be connected to the presence of an interacting network between the C-terminus and the rest of the protein stabilising the *closed* form. It may also imply that the presence of L-Kyn strengthens this network, as discussed in section 1.6.3.

Figure 3.25 is of substrate-bound molecule A; it shows the residues identified in chapter 1 as important in forming an interactive network.

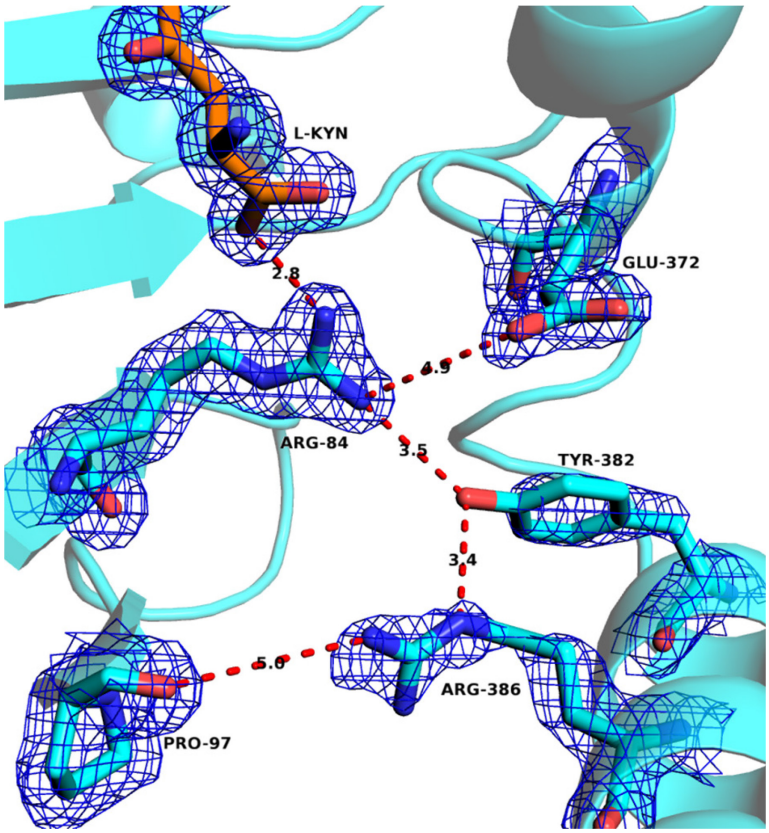

Figure 3.25 Crystal structure of putative C-terminus interactive network of substrate-bound dm2, molecule A in the closed conformation. 2Fo-Fc map is shown in blue (contoured at 0.34 eÅ^{-3}, 1.0 σ) Distances between atoms are in Å and shown as red dashed lines.

The interactions between the residues shown are significantly decreased when compared to the initial models discussed in section 1.6.3. Only residues R84, Y382 and R386 are close enough to interact, and these distances would only imply very weak interactions. The electron density surrounding residues Y382 and R386 is also quite sparse. The average B-factors of the residues are 64.2 Å^2 and 49.5 Å^2, respectively; much higher than the average B-factor for the whole molecule of 35.1 Å^2. This is interpreted as indicating that these residues in particular are very mobile. This is also seen in the other molecules of dm2, as shown in Figure 3.26.

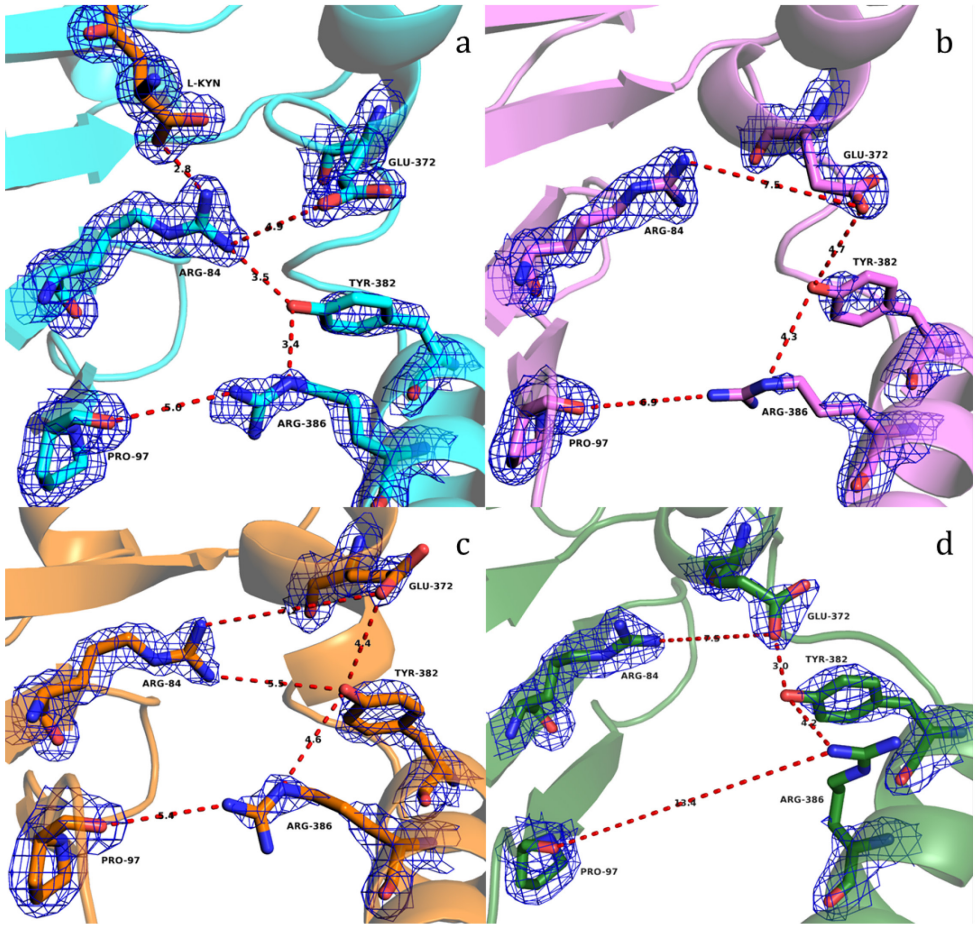

Figure 3.26 Crystal structure showing residues involved in the putative interacting network of the C-terminus of dm2 **a.** Closed molecule of L-Kyn bound dm2 **b.** Open molecule of L-Kyn bound dm2 **c.** Closed molecule of L-Kyn free dm2 **d.** Open molecule of L-Kyn free dm2. 2Fo-Fc maps are shown in blue (L-Kyn bound molecules contoured at 0.34 eÅ^{-3}, 1.0 σ, L-Kyn free molecules contoured at 0.32 eÅ^{-3}, 1.0 σ). Distances between atoms are in Å and shown as red dashed lines.

From all the residues shown in Figure 3.26 only one interatomic distance is seen that would imply interaction. That is between E372 and Y382 in the *open* molecule B of substrate free dm2. The distances measured between the other residues are too large for them to be interacting. Despite the high resolution and low overall B-factor values of each molecule the electron density surrounding these residues, in particular R386 and Y382, is very limited. This implies that the mobility of these residues is very high.

While this does mean that the exact positioning of the residues is difficult to pinpoint, the high mobility itself shows that there is no obvious network of interactions between the residues. If strong interactions were present the residues would be expected to be held in place by them, and the B-factors considerably lower. This matches with the B-factors exhibited in Table 3.10, as distances between these residues increase the B-factors increase accordingly. This reasoning is somewhat indirect and the questions surrounding the mechanistic importance of the C-terminus would benefit from further directed study.

Another confusing factor in the analysis of the movement of the C-terminus is the surface calculation data of the active site and channel opening. The expected pattern would be the *closed* molecule having a lower void volume and surface area. The results presented in this chapter counter this expectation, particularly the results of the substrate-free model, seen in Table 3.8. Further surface calculations are required of the other *Pf*KMO variants before a hypothesis can be suggested. These are discussed in chapter 5.

3.5. Summary

Throughout this chapter a new crystal form of dm2, grown using PEG 8K conditions, has been examined. Thanks to a consistent space group between substrate-free and substrate-bound crystals comparisons between the two have been made easier. Interactions between L-Kyn and the active site have been recognised and now require further probing to establish their importance. Movement of the C-terminus has been shown to be essentially unaffected by presence or absence of substrate and both *open* and *closed* conformers of dm2 are present in each crystal. It has also been shown that crystal contacts do not have an effect upon the position of the C-terminus. The surface area and void volume of the dm2 active site does not appear to show any correlation between the *closed* and *open* conformers. Movement in residues 240-290 has been observed for the first time and is posited to have a connection to the *open* and *closed* conformations. Very little evidence has been seen of an interactive network within the C-terminus. In light of all these finding the role of the C-terminus, in terms of dm2's mechanism, is still to be established. The importance of the residues within this putative interactive network is also in need of clarification.

Chapter 4: Binding residues

Certain residues within the active site of dm2 have been previously identified by M. Wilkinson as important for the binding of L-Kyn[143]. These residues have been described in Chapters 1 and 3 and their importance has been identified through structural analysis of dm2. The interactions were then investigated by H. Bell by creating a number of substitution mutations within the active site of dm2, followed by kinetic analysis of these mutants[156]. The Y404F mutation was introduced as it was a conservative substitution and allowed for the investigation of the H-bond between Y404 and L-Kyn. In the original model of dm2 the distance between substrate and Y404 had appeared to be too great to allow an H-bond. The removal of H-bond capability would determine the effect that Y404 interactions have upon substrate binding. The R84K mutation was also conservative given the similar size and positive charge on both residues. This would allow for investigation of the importance of interaction between substrate and R84. The contribution of N369 in substrate binding was also investigated with an N369S mutation. An H320F mutant was also produced since the mutation gives the protein an entirely analogous active site (in terms of sequence and likely structure) to that of hKMO, as seen in Figure 1.10. This allows for a closer structural representation of hKMO than previously seen.

The Michaelis-Menten kinetic data for the mutant forms of dm2 are shown in Table 4.1.

Enzyme	K_M (μM)	k_{cat} (s^{-1})	k_{cat}/K_M ($\mu M^{-1} s^{-1}$)
dm2	8.8 ± 1.5	8.9 ± 1.1	1.01
Y404F	96.5 ± 4.0	6.2 ± 0.1	0.06
R84K	0.0	0.0	0.0
N369S	27.7 ± 3.3	4.5 ± 0.2	0.16
H320F	18.0 ± 1.2	7.1 ± 0.2	0.39

Table 4.1 Summary of Michaelis-Menten kinetic data for mutant forms of dm2. R84K exhibits no activity and is reported as zero[156].

The kinetic data for the mutant forms of dm2 described in Table 4.1 requires structural rationalisation. While structures of the mutants have already been solved the crystals were formed using the old PEG 4K conditions. This meant that substrate-free and substrate-bound forms were in different space groups making comparisons more difficult. The resolution of these structures, especially the substrate-bound forms, was very low. The highest resolution for the substrate-bound models never exceeded 3.9 Å and the structure of a substrate-bound R84K mutant was never obtained. Using the new PEG 8K conditions the resolutions of these structures were greatly improved. This chapter presents the structures of the

dm2 mutants R84K, Y404F and H320F. Y404F has a ~ 30% reduction of k_{cat} value but has a K_M 10-fold greater than that of dm2. This implies Y404 is an important residue involved in substrate binding. R84K exhibited no activity whatsoever, implying it also is involved in substrate binding or mechanism. The change in kinetic parameters for N369S was not considered great enough to be a priority for structural rationalisation. The slight increase in K_M may be accounted for in the size difference between Asn and Ser. The H-bonding group of Ser may be too far from L-Kyn to form a meaningful interaction in N369S. H320F, on the other hand, was investigated given its sequence similarity with hKMO, as previously mentioned.

4.1. Substrate bound R84K dm2

4.1.1. Crystal growth and data collection

R84K dm2 protein was expressed and purified as described in Chapter 2 using the plasmid created by H. Bell[156]. The protein was concentrated to 300 µM, confirmed by UV-Vis spectrophotometry as shown in Figure 4.1.

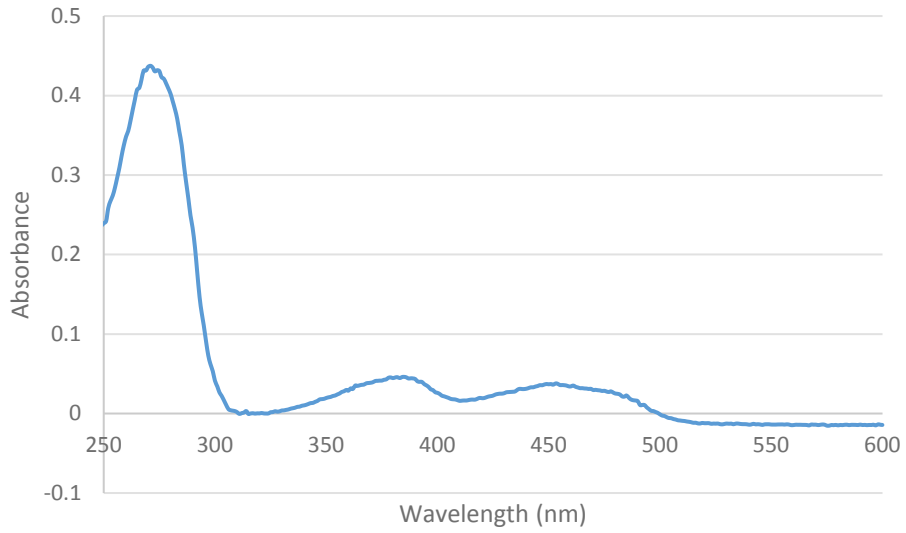

Figure 4.1 UV-visible spectrum of 1:100 dilution of R84K. Concentration calculated from absorbance of 0.037 at 450 nm with ε_{450} of 12,300 M^{-1} cm^{-1}.

The new PEG 8K conditions were used to set up hanging drop trays as described in Chapter 2. Two trays were set up, one co-crystallised with 1 mM L-Kyn, the other without L-Kyn. Within 24 hours a number of large crystals had formed. Selected crystals were then soaked in 1 µL drops comprising mother liquor stabilised with extra PEG 8K and 1 mM L-Kyn. Crystals were then left for a further 24 hours. Individual crystals were flash cooled in liquid N_2 and sent to the Diamond Light Source. Using

beamline i04-1 data were collected. All crystals exhibited $P12_11$ spacegroup with similar parameters to those seen in Table 3.4.

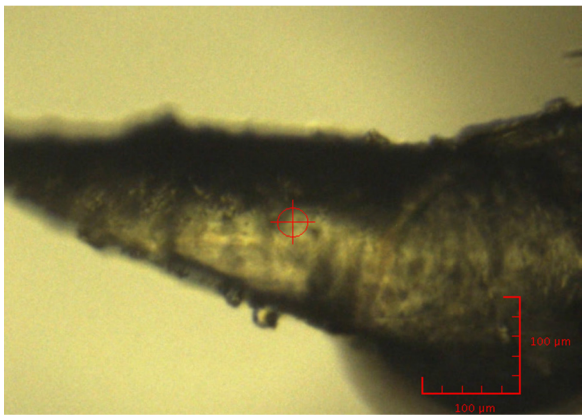

Figure 4.2 Mounted crystal of substrate-exposed R84K used for data collection.

From the crystal shown in Figure 4.2 1800 images were collected with beamline transmission 18.75 %, exposure of 0.05 s and oscillation of 0.15°. The statistics shown in Table 4.2 were calculated using the programs *iMosflm* and *Scala*[159, 160].

Resolution (Å)	R_{merge}	Mean (I)/σ(I)	Completeness (%)	Multiplicity
68.20-1.70	0.046	10.9	94.5	2.9

Table 4.2 Data collection statistics for a substrate-bound R84K crystal. All calculated from overall data, for full data collection statistics see Appendix 7.1.2.

Refinement and model editing were performed as for dm2, described in section 2.8. The structure was solved with two molecules in the ASU and a solvent content of 49.4%. An overall image of the two molecules in the ASU can be seen in Figure 4.3.

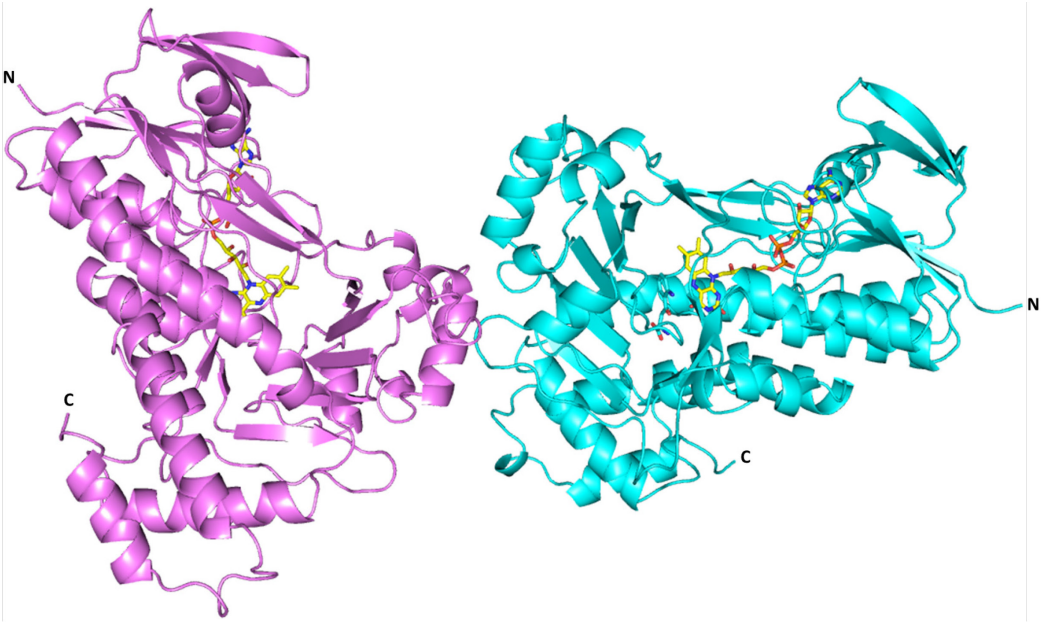

Figure 4.3 Crystal structure of substrate-exposed R84K showing both molecules in the ASU. Molecule A is coloured cyan, molecule B is coloured violet.

4.1.2. Discussion of substrate-bound R84K dm2 structure

The model showed both molecules with residues 6-457 and contained 659 waters in the ASU. The final model was refined to give R and R_{free} values of 0.1655 and 0.1966, respectively, to a resolution of 1.70 Å. The full statistics can be seen in Appendix 7.1.2.

When comparing the two molecules within the ASU a difference in structure is observed in the C-terminus and residues 240-290, as shown in Figure 4.4. These follow the movements discussed in sections 1.6.3 and section 3.2.5, respectively.

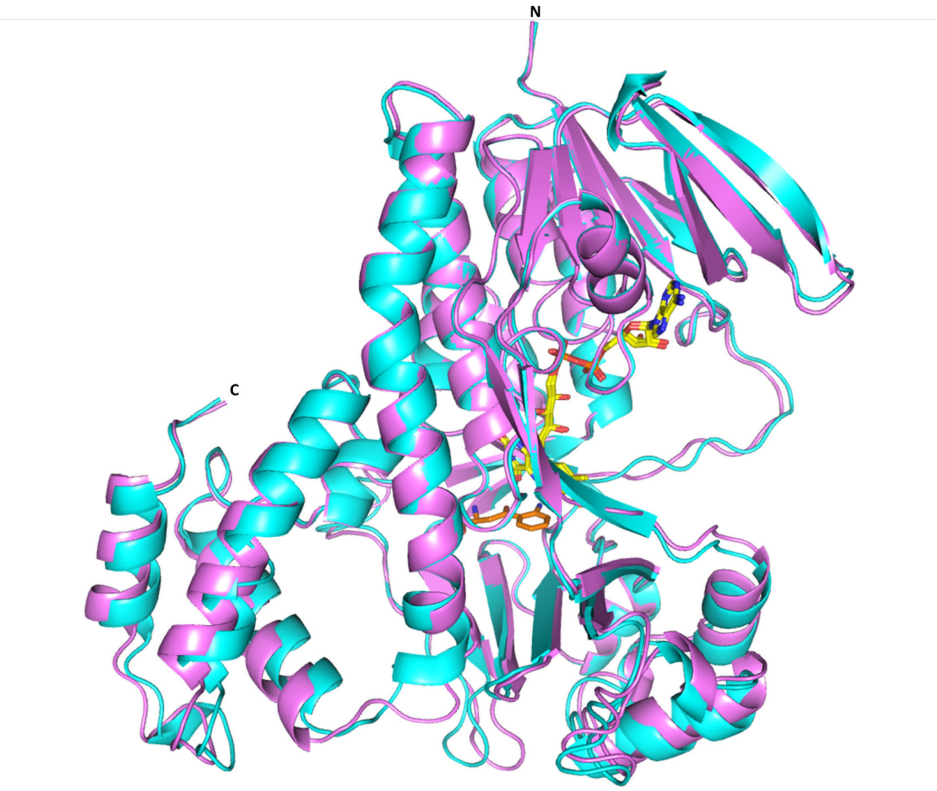

Figure 4.4 Crystal structure of substrate-bound R84K showing both molecules overlaid with one another. Molecule A is coloured cyan, molecule B is coloured violet. L-Kyn is coloured orange. RMSD = 0.937 Å from 445 paired Cα atoms.

Molecule	Mean B-factor of C-terminus/Mean B-factor of structure
Closed	1.1
Open	1.6

Table 4.3 Average B-factor for residues 370-400 of substrate-bound R84K divided by average B-factor of overall structure.

The overall structures of R84K and dm2 are very similar, with an RMSD of less than an ångström. The absence of R84 does not appear to affect the movement of the C-terminus given that the positional shift is similar and the relationship between B-factors of the two molecules are similar to dm2, as shown in Table 4.3. K84 and its dm2 counterpart R84 are compared in Figure 4.5.

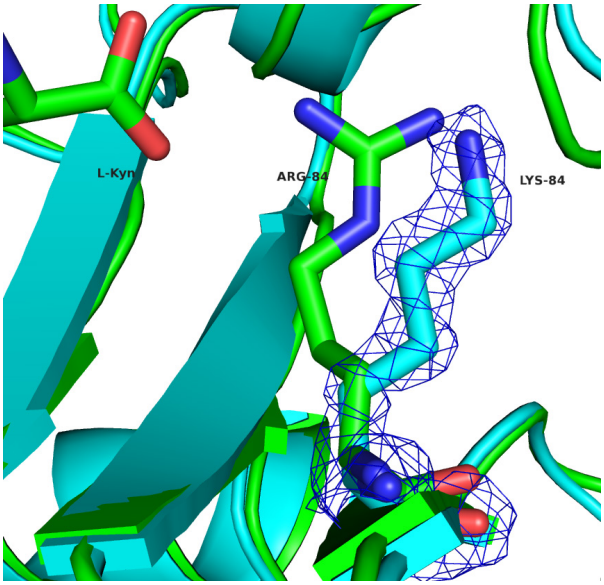

Figure 4.5 Crystal structures of substrate-bound R84K and substrate-bound dm2 overlaid, showing residue 84. RMSD = 0.573 Å from 447 aligned Cα atoms. R84K is shown in cyan, dm2 is shown in green, 2Fo-Fc map is shown in blue (contoured at 0.47 eÅ^{-3}, 1.50 σ).

Firstly there is a difference in structure for lysine only allows monodentate interaction, unlike the guanidinium of arginine. R84 has previously shown bidentate interaction with L-Kyn in the original model shown in Figure 1.25. Secondly the positioning of K84 points the residue away from the substrate, meaning there can be no interaction between L-Kyn's acid group and the lysine's amine. The resulting distance between substrate and K84 is shown in Figure 4.6.

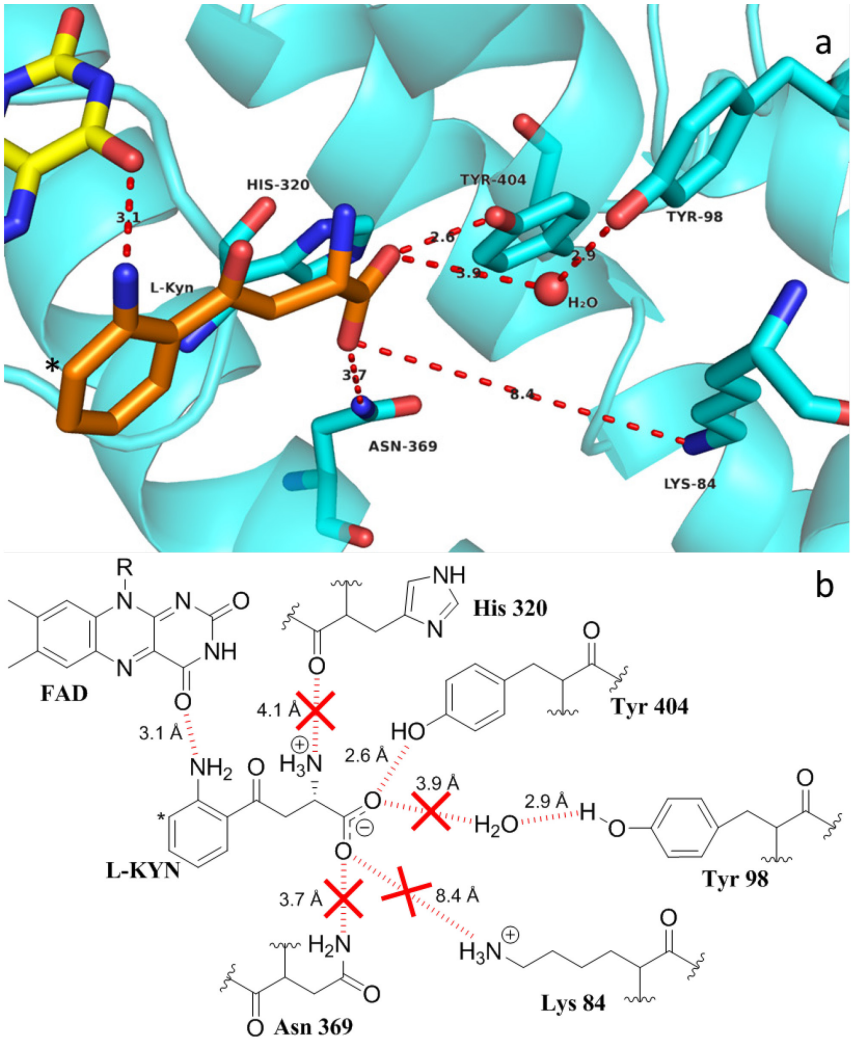

Figure 4.6 a. Crystal structure of active site of R84K bound with L-Kyn, molecule A. **b.** Chemdraw schematic of R84K bound with L-Kyn. Distances between atoms are in Å and shown as red dashed lines. The site of oxidation is denoted with an asterisk (*).

There are a number of differences between the substrate binding sites of R84K, shown here, and dm2, shown in Figure 3.14. First of all, the distance from residue 84 to substrate is now over 8 Å, which implies no interaction. In addition to this, N369 is also too far from the substrate to form a significant interaction. Y404 is within H-bonding distance but now interacts with the acid group of L-Kyn, rather than the amine. The acid group of L-Kyn has rotated 90° due to the absence of R84 and in order to form a bond with Y404. Y98 still forms an H-bond to a water molecule but the distance and direction of the L-Kyn acid group relative to this water rules out any meaningful interaction.

The reduced interaction with L-Kyn implies a reduced binding affinity for the enzyme; this in turn results in a lower occupancy of substrate. The presence of substrate in R84K is much less than in dm2, as shown in Figure 4.7.

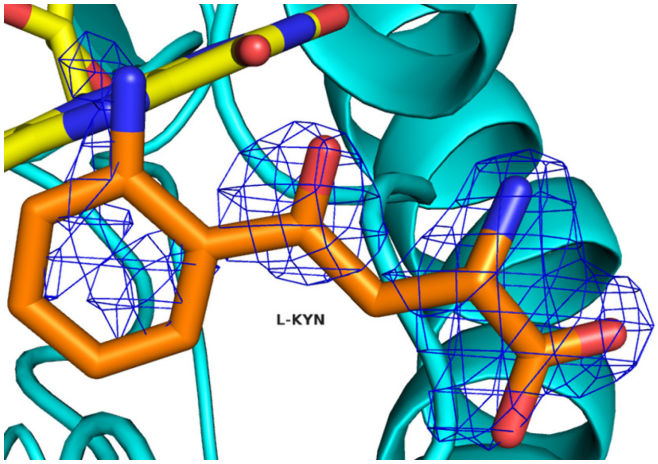

Figure 4.7 Crystal structure of substrate-bound R84K showing L-Kyn. 2Fo-Fc map is shown in blue (contoured at 0.31 eÅ^{-3}, 1.00 σ)

The electron density surrounding L-Kyn is sparse when contoured to the same level as dm2, the contouring was therefore decreased. This shows the map surrounding the atoms with highest electron density, as would be expected. The electron density surrounding the amino acid group in particular supports the presence and positioning of L-Kyn rather than only water molecules being present within the active site. L-Kyn is modelled within the structure with an occupancy of 100 %; this also supports the conclusion that L-Kyn is bound, if somewhat weakly, within the molecule. This is discussed further in section 4.2.2.

The interruption in interactions between L-Kyn and certain binding residues cannot in itself explain the total loss of activity seen in R84K. However, when comparing the substrate molecules of R84K and dm2 a reason can be suggested.

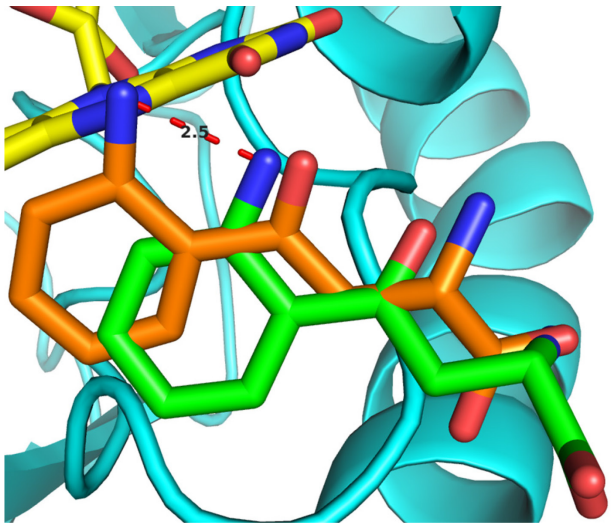

Figure 4.8 Crystal structure of R84K showing L-Kyn overlaid with L-Kyn of dm2. R84K is shown in orange, dm2 is shown in green. Distances between atoms are in Å and shown as red dashed lines. RMSD = 0.573 Å from 447 aligned Cα atoms.

When compared in Figure 4.8 the relative position of the modelled L-Kyn in dm2 and R84K has changed. The molecule has moved 2.5 Å away from N369 and Y404. dm2 shows all of L-Kyn below the plane of the isoalloxazine of FAD whereas in R84K the aromatic amine of L-Kyn is in line with the plane. This change in position is enough to rationalise the inactivity of the enzyme, since its position in relation to the FAD may no longer allow for a reaction with the hydroperoxide intermediate to occur.

4.2. Substrate-free R84K dm2

4.2.1. Crystal growth and data collection

Using the protein prepared in section 4.1 new crystal trays were prepared using the same method as described in section 4.1.1 with no exposure to L-Kyn. These substrate-free crystals were sent to Diamond Light Source. Using beamline i03 data were collected.

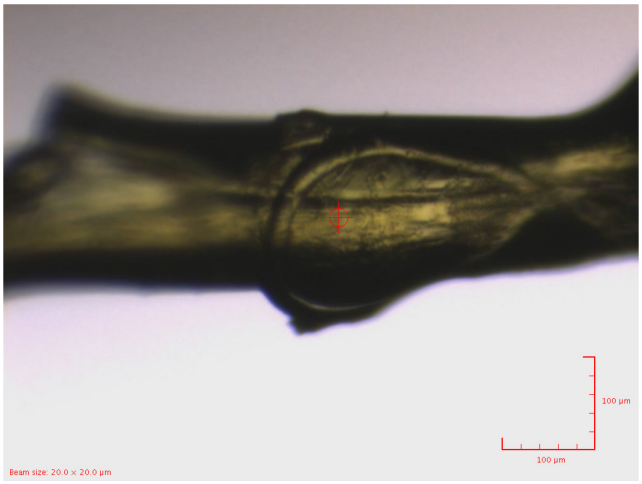

Figure 4.9 Mounted crystal of substrate-free R84K used for data collection.

From the crystal shown in Figure 4.9 2000 images were collected with beamline transmission 100 %, exposure of 0.1 s and oscillation of 0.1°. The statistics shown in Table 4.4 were calculated using the programs *iMosflm* and *Scala*[159, 160].

Resolution (Å)	R_{merge}	Mean (I)/σ(I)	Completeness (%)	Multiplicity
67.95-1.80	0.113	4.6	98.3	2.7

Table 4.4 Data collection statistics for a substrate-free R84K crystal. All calculated from overall data, for full data collection statistics see Appendix 7.1.2.

Refinement and model editing were performed as for dm2, described in section 2.8. The structure was solved with 2 molecules containing residues in the ASU and a solvent content of 49.2 %. Each molecule

contained residues 6-457 and a total of 614 waters were modelled in the ASU. The final model was refined to give R and R_{free} values of 0.1879 and 0.2283, respectively, to a resolution of 1.80 Å. The full statistics can be seen in Appendix 7.1.2.

4.2.2. Discussion of substrate-free R84K dm2 structure

The overall structure of substrate-free R84K is similar to the substrate-bound form of R84K. There is however a difference in the shift of the C-terminus, as seen in Figure 4.10.

Figure 4.10 Crystal structure of substrate-free R84K, showing both molecules overlaid with one another. Molecule A is coloured orange, molecule B is coloured green. RMSD = 0.582 Å from 447 aligned Cα atoms.

Molecule	Mean B-factor of C-terminus/Mean B-factor of structure
A	1.5
B	1.8

Table 4.5 Average B-factor for residues 370-400 of substrate-free R84K divided by average B-factor of overall structure.

There is very little movement of the C-terminus seen between the two molecules. The average B-factor of the residues within the C-terminus are shown in Table 4.5. Given the lack of a clear *open* or *closed* molecule it is difficult to claim the numbers follow the trend of the closed molecule having the lower B-factor. When compared to the substrate-bound molecules the substrate-free C-termini are in between the *closed* and *open* position.

The decreased movement of the C-terminus was seen in other substrate-free R84K crystals. This is therefore not an occurence in one crystal. Some crystals showed the two molecules to be similar enough to only assign one molecule within the ASU, with the unit cell halved in length.

Interestingly the movement in residues 240-290 is also minimilised. This supports the link between the two sections; since the C-terminus is not moving, residues 240-290 no longer need to accommodate said movement.

The position of K84 in all four molecules is shown in Figure 4.11.

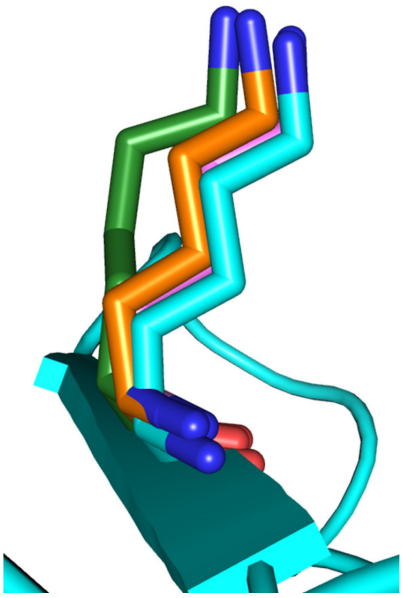

Figure 4.11 Crystal structure of R84K molecules, showing residue K84 from both molecules of substrate-bound and substrate-free structures. RMSD values were generated by aligning each molecule to the "closed" substrate-bound molecule (cyan). RMSD values of the overlays range from 0.729 – 0.965 Å, from at least 439 aligned Cα atoms in each case.

The active sites within all four molecules of both the substrate-bound and substrate-free crystals are very similarly positioned. Much like dm2 the binding residues are not seemingly affected by the presence or absence of L-Kyn.

In order to provide further evidence of L-Kyn occupancy in the substrate-exposed structure the Fo-Fc maps of the substrate-exposed structure and the substrate-free structure were compared. This is to counter the possibility that the electron density is only water molecules. L-Kyn and water molecules were deleted from the active sites of the structures and these models were refined using *Refmac5*[164]. The results are shown in Figure 4.12.

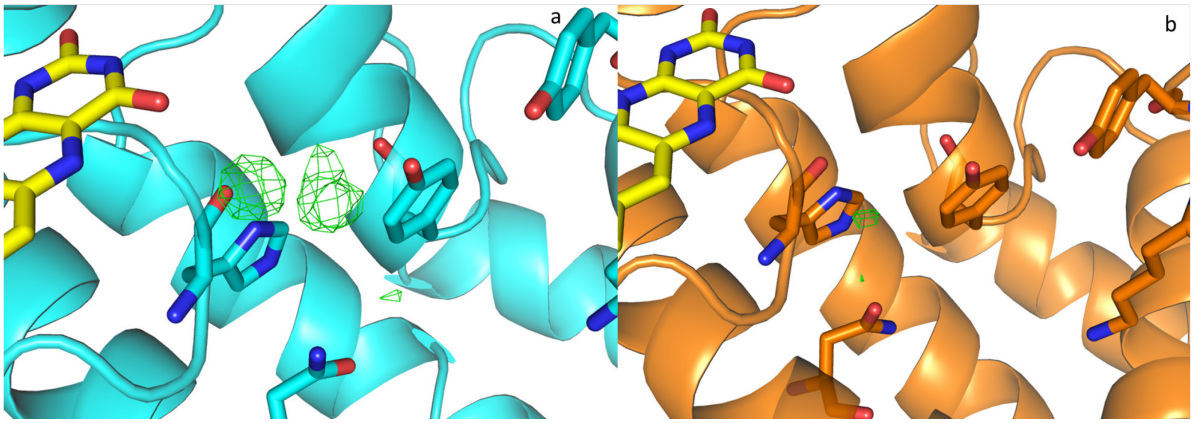

Figure 4.12 Crystal structures of R84K showing the active site. **a.** Substrate-exposed R84K with Fo-Fc map (contoured at 0.31 eÅ^{-3}, 3.00 σ) **b.** Substrate-free R84K with Fo-Fc map (contoured at 0.31 eÅ^{-3}, 3.00 σ).

The Fo-Fc maps shown in Figure 4.12 show a clear difference in occupancy levels between the two structures. This supports the presence of the L-Kyn in the substrate-exposed structure, if the active site had only water molecules present we would expect to see similar peaks in the Fo-Fc map of the substrate-free structure.

There is also the possibility that the electron density modelled in Figure 4.7 is caused by acetate or glycerol, both of which are present in the crystallising well solution. These small molecules are used in the well solution for both substrate-exposed and substrate-free crystals. One would, therefore, expect to see the same levels of acetate and glycerol in both crystal structures. There is, however, an absence of any significant electron density within the substrate-free protein in Figure 4.12. This indicates the difference in electron density between the two structures at the active site is caused by weakly bound L-Kyn.

The use of polder maps was also implemented to assess the presence of L-Kyn. The substrate molecule was selected for omission and a polder map was created using *Phenix*[170, 171]. The three maps were then compared using correlation coefficients; see section 2.8.2 for details.

Maps being compared	Local correlation coefficient	Peak correlation coefficient
1, 2	0.69	0.76
1, 3	0.81	0.76
2, 3	0.68	0.72

Table 4.6 Correlation coefficients of polder map calculations for substrate-bound R84K.

The correlation is greatest between maps 1 and 3 the polder map, supporting the presence of L-Kyn. However, the similar values of peak correlation coefficients require consideration and imply that these results are insufficient to decide whether the density belongs to ligand or bulk solvent[171]. These findings,

if used in isolation would not allow for strong conclusions. The results can, however, be added to the other supporting factors, such as the Fo-Fc maps and the high ligand occupancy. The use of all these data bolsters the argument for the presence of a weakly bound substrate but cannot definitively confirm it.

The findings of this section matches with previous work by Mole *et al.,* Amaral *et al.* and Bell in highlighting the importance of residue R84[11, 140, 156]. This brings a structural rationalisation to the inactivity seen in R84K *Pf*KMO and the drop in activity in the equivalent R83 mutants of *Sc*KMO.

4.3. Substrate-bound Y404F dm2

4.3.1. Crystal growth and data collection

Y404F dm2 protein was expressed and purified as described in Chapter 2 using the plasmid created by H. Bell[156]. The protein was concentrated to 350 µM confirmed by UV-Vis spectrophotometry, as shown in Figure 4.13.

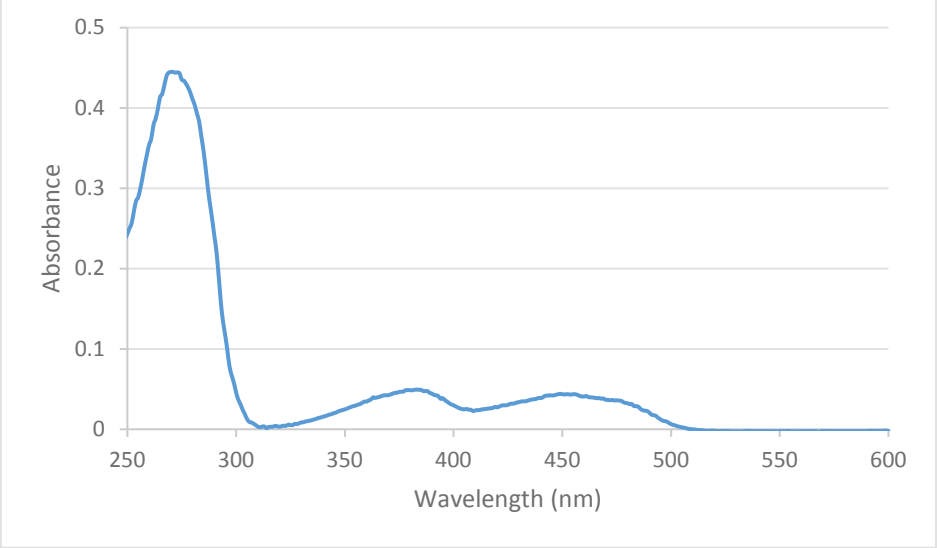

Figure 4.13 UV-visible spectrum of 1:100 dilution of Y404F. Concentration calculated from absorbance of 0.043 at 450 nm with ε_{450} of 12,300 M^{-1} cm^{-1}.

The new PEG 8K conditions were used to set up hanging drop trays as described in Chapter 2. Two trays were set up, one co-crystallised with 1 mM L-Kyn, the other without L-Kyn. Crystals formed within 24 hours. Selected crystals were then soaked in 1 µL drops comprising of mother liquor stabilised with extra PEG 8K and 1 mM L-Kyn. After a further 24 hours individual crystals were flash cooled in liquid N_2 and sent to the Diamond light source. Data were collected using beamline i04-1. All crystals exhibited $P12_11$ spacegroup with dimensions similar to those stated in Table 3.4.

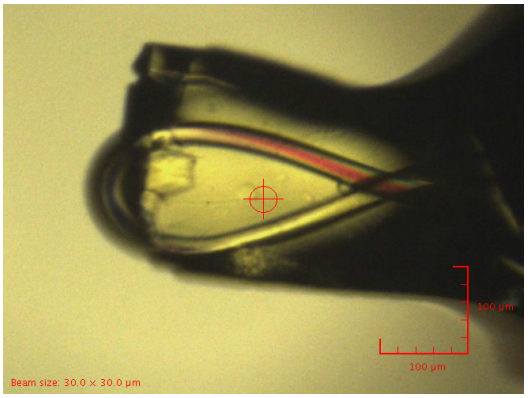

Figure 4.14 Mounted crystal of substrate-exposed Y404F used for data collection.

From the crystal shown in Figure 4.14 1800 images were collected with beamline transmission 18.75 %, exposure of 0.05 s and oscillation of 0.15°. The statistics shown in Table 4.7 were calculated using the programs *iMosflm* and *Scala*[159, 160].

Resolution (Å)	R_{merge}	Mean (I)/σ(I)	Completeness (%)	Multiplicity
68.20-1.80	0.069	8.4	96.0	3.1

Table 4.7 Data collection statistics for a substrate-bound Y404F crystal. All calculated from overall data, for full data collection statistics see Appendix 7.1.3.

Refinement and model editing were performed as for dm2, described in section 2.8. The structure was solved with 2 molecules in the ASU and a solvent content of 49.8 %.

4.3.2. Discussion of substrate-bound Y404F dm2 structure

Each molecule contained residues 7-457 and a total of 847 waters were modelled in the ASU. The final model was refined to give R and R_{free} values of 0.1667 and 0.2023, respectively, to a resolution of 1.80 Å. The full statistics can be seen in Appendix 7.1.3. The overall structure of the molecule is very similar to dm2. Once again the movement in the C-terminus is seen, as shown in Figure 4.15. In this structure however there is very little movement in residues 240-290. This contradicts the theory that the two sections are connected, or at least that the connection is particularly strong. Given this lack of movement it could be argued that residues 240-290 are just highly mobile and subject to thermal motion.

Figure 4.15 Crystal structure of substrate-bound Y404F, showing both molecules overlaid with one another. Molecule A is coloured in cyan, molecule B is coloured violet. L-Kyn is coloured orange. RMSD = 0.995 Å from 445 aligned Cα atoms.

Molecule	Mean B-factor of C-terminus/Mean B-factor of structure
Closed	1.3
Open	1.3

Table 4.8 Average B-factor for residues 370-400 of substrate-bound Y404F divided by average B-factor of overall structure.

The B-factors shown in Table 4.8 show the same value for the *closed* molecule and the *open* molecule. This goes against the previous findings. While many datasets for the Y404F protein matched the pattern of the *closed* molecule showing lower B-factors they did not show the same level of L-Kyn occupancy. Therefore this model was selected. The B-factors shown in Table 4.8 are also closer than any other structure, meaning the two molecules are very similar in mobility.

As shown in Figure 4.16 the positioning of F404 is very similar to that of Y404 in dm2. The similarity between the two amino acids makes it difficult to ensure presence of a phenylalanine residue over a tyrosine. When a tyrosine residue is used in the model however the Fo-Fc map clearly shows the presence of the hydroxyl is not appropriate, as shown in Figure 4.16.

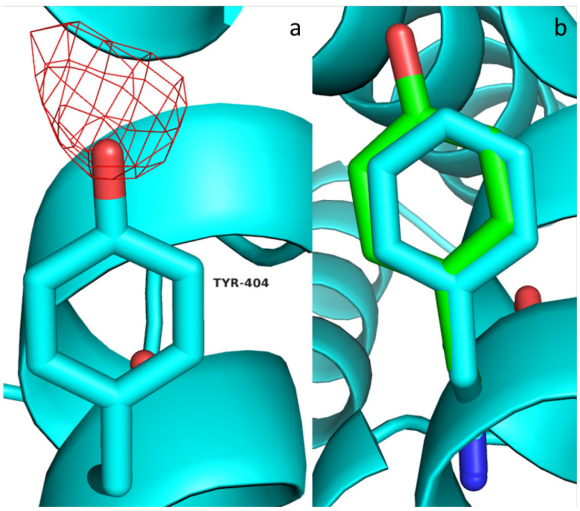

Figure 4.16 a. Crystal structure of substrate-bound Y404F, modelled with Tyr as residue 404. Fo-Fc map is shown in red (contoured as -0.32 eÅ^{-3}, 3.25 σ). **b.** Crystal structure of substrate-bound Y404F overlay with dm2 showing residues 404. Y404F coloured cyan, dm2 coloured green. RMSD = 0.419 Å from 446 aligned Cα atoms.

Only the absence of the hydroxyl affects a change on the binding of the substrate. Once again the occupancy of L-Kyn is much lower than seen in dm2. In order to model L-Kyn the contouring of electron density was decreased, as shown in Figure 4.17. This is discussed further in section 4.4.2.

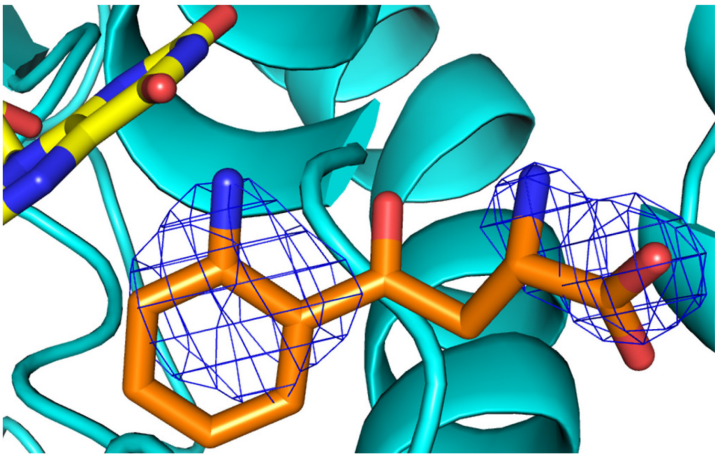

Figure 4.17 Crystal structure of substrate-bound Y404F, showing L-Kyn. 2Fo-Fc map shown in blue (contoured at 0.24 eÅ-3, 0.75 σ).

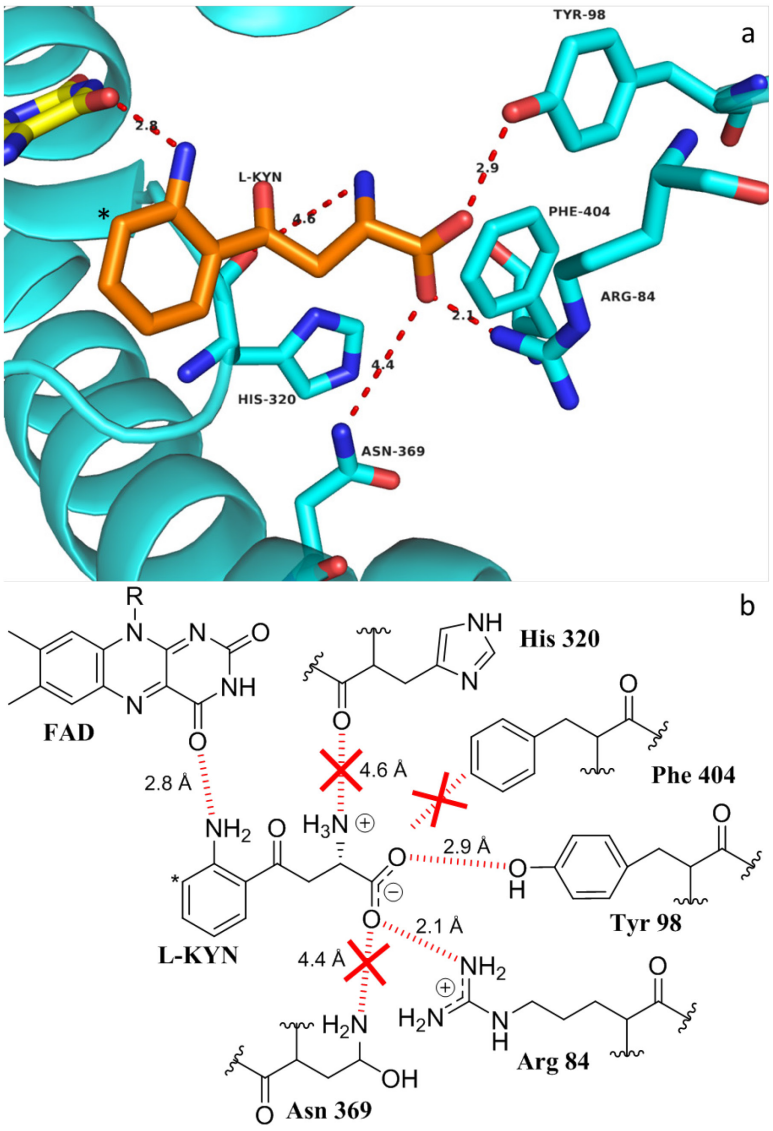

Figure 4.18 a. Crystal structure of active site of Y404F bound with L-Kyn, molecule A. **b.** Chemdraw schematic of Y404F bound with L-Kyn. Distances between atoms are in Å and shown as red dashed lines. The site of oxidation is denoted with an asterisk (*).

As seen in Figure 4.18, L-Kyn shows similar positioning to that seen in dm2. The acid group positioning implies a direct H-bond to Y98, unlike in dm2. The acid remains close enough to R84 to form an electrostatic interaction, as exhibited in dm2. The distance from the acid to N369 is now too large to form any meaningful interaction. As a result of the substitution the amine has not only lost the Y404 interaction but also the H320 interaction it showed in dm2. This could allow for the slight movement of L-Kyn resulting in this new distribution of binding interactions. Despite this the overall positioning of L-Kyn is still close enough to allow interaction between it and the FAD. This means catalysis can still occur, unlike in R84K. Where dm2 exhibited 6 meaningful interactions between substrate and binding site Y404F only shows 3. This bonding form is clearly less effective than that seen in dm2, as shown by the large increase in K_M for Y404F.

4.4. Substrate-free Y404F dm2

4.4.1. Crystal growth and data collection

Using the protein prepared in section 4.3 new crystal trays were prepared using the same method as described in section 4.3.1 with no exposure to L-Kyn. These substrate-free crystals were sent to Diamond Light Source. Using beamline i03 data were collected.

Figure 4.19 Mounted crystal of substrate-free Y404F used for data collection.

From the crystal shown in Figure 4.19 1500 images were collected with beamline transmission 100 %, exposure of 0.1 s and oscillation of 0.15°. The statistics shown in Table 4.9 were calculated using the programs *iMosflm* and *Scala*[159, 160].

Resolution (Å)	R_{merge}	Mean (I)/σ(I)	Completeness (%)	Multiplicity
67.90-1.65	0.070	8.7	99.8	3.7

Table 4.9 Data collection statistics for a substrate-free Y404F crystal. All calculated from overall data, for full data collection statistics see Appendix 7.1.3.

Refinement and model editing were performed as for dm2, described in section 2.8. The structure was solved with 2 molecules in the ASU and a solvent content of 48.7 %.

4.4.2. Discussion of substrate-free Y404F dm2 structure

Each molecule contained residues 6-459 and a total of 605 waters were modelled in the ASU. The final model was refined to give R and R_{free} values of 0.1748 and 0.2099, respectively, to a resolution of 1.65 Å. The full statistics can be seen in Appendix 7.1.3. The movement in the C-terminus was also exhibited between the two molecules within the ASU, as shown in Figure 4.20. Once again there is no movement observed in residues 240-290, implying no strong relationship between these residues and the C-

terminus exists. The possibility that residue Y404 is important to this movement is very remote; given Y404 is positioned very far from residues 240-290.

Figure 4.20 Crystal structure of substrate-free Y404F, showing both molecules overlaid with one another. Molecule A is coloured in orange, molecule B is coloured green. RMSD = 0.732 Å from 456 aligned Cα atoms.

Molecule	Mean B-factor of C-terminus/Mean B-factor of structure
Closed	1.4
Open	1.8

Table 4.10 Average B-factor for residues 370-400 of substrate-free Y404F divided by average B-factor of overall structure.

The B-factors, shown in Table 4.10, return to the expected relationship of the closed molecule having a lower value. The expected movement in C-terminus is also observed between the two molecules. This, however, was not the case in all the substrate free Y404F crystals. Much like the substrate free R84K crystals, some were capable of being solved with only one molecule within the ASU. Meaning the movement of C-terminus was small enough to assign to two molecules as crystallographically identical.

The positioning of F404 is very similar between the two substrate-free molecules, as well as to the substrate-bound model. An overlay of the four molecules is shown in Figure 4.21.

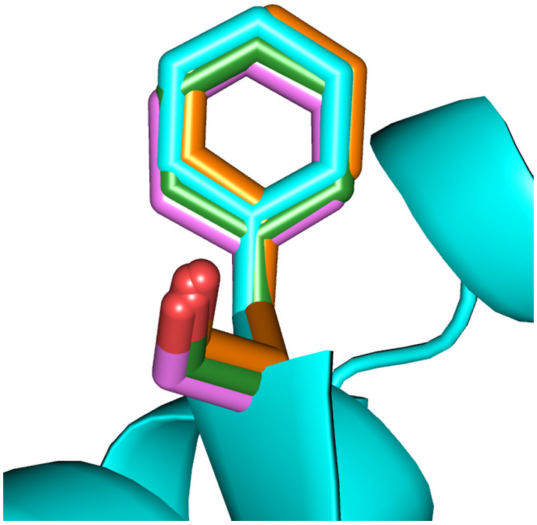

Figure 4.21 Crystal structure of Y404F molecules, showing residue F404 from both molecules of substrate-bound and substrate-free structures. RMSD values were generated by aligning each molecule to the "closed" substrate-bound molecule (cyan). RMSD values of the overlays range from 0.578 – 1.037 Å, from at least 445 aligned Cα atoms in each case.

This shows that, despite the unexpected behaviour in terms of the C-terminus, the F404 residue is totally unaffected by presence or absence of L-Kyn.

Much like R84K the Fo-Fc maps of the substrate-exposed and substrate-free structures were compared. Firstly the substrate and any water molecules were removed from the models, these were then refined using *Refmac5*[164]. The results are shown in Figure 4.22.

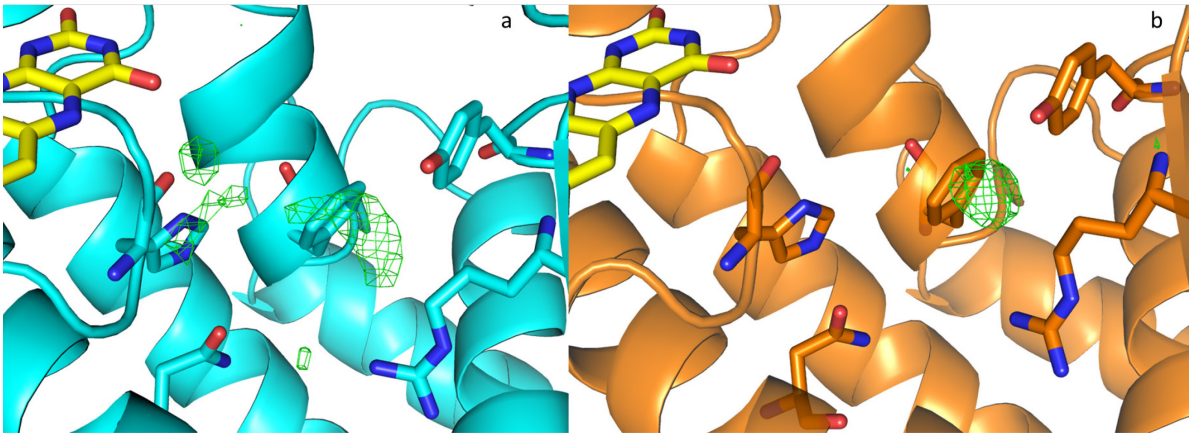

Figure 4.22 Crystal structures of Y404F showing the active site. **a.** Substrate-exposed Y404F with Fo-Fc map (contoured at 0.29 eÅ^{-3}, 2.95 σ). **b.** Substrate-free Y404F with Fo-Fc map (contoured at 0.31 eÅ^{-3}, 2.96 σ).

There are once again differences between the substrate-exposed and substrate-free Fo-Fc maps. These are not quite as marked as in the R84K structures because of the green peak next to Y98 in both structures. An absence of electron density correlating to acetate or glycerol in the substrate-free

molecule reduces the likelihood that these molecules correspond to electron density in the substrate-bound structure. In Figure 4.22 b. the peak is spherical and can be attributed to the water molecule seen in the dm2 structures in Figure 3.14 and Figure 3.23, rather than acetate or glycerol. The occupancy of L-Kyn is so low in the substrate-exposed Y404F; many molecules within the crystal will only have this water molecule present, while others will have L-Kyn bound. This makes the modelling of L-Kyn more difficult. However, one can still see differences between these two peaks. There are also more peaks in the Fo-Fc map of the substrate-exposed model. This supports the case for L-Kyn presence, albeit at low occupancy.

The use of polder maps was again implemented to assess the presence of L-Kyn. The substrate molecule was selected for omission; a polder map was then created using *Phenix*[170, 171]. The three maps were then compared using correlation coefficients; see section 2.8.2 for details.

Maps being compared	Local correlation coefficient	Peak correlation coefficient
1, 2	0.66	0.73
1, 3	0.83	0.80
2, 3	0.55	0.58

Table 4.11 Correlation coefficients of polder map calculations for substrate-bound Y404F.

Given the correlation is greatest between maps 1 and 3 the polder map supports the presence of L-Kyn. Maps 1 and 3 also show the higher peak correlation coefficient, increasing the likelihood of substrate presence. Compared to R84K these findings are more definitive. When used in consideration of the other supporting factors, such as the Fo-Fc maps and the high ligand occupancy, the data more strongly reinforce the argument for the presence of a weakly bound substrate.

This section has shown the importance of the interaction between Y404 and L-Kyn. When this interaction breaks down the binding ability of the protein is affected, as exhibited in the kinetic and structural data, with reduced binding interactions leading to a higher K_M and lower k_{cat}.

4.5. Substrate-bound H320F dm2

4.5.1. Crystal growth and data collection

H320F dm2 protein was expressed and purified as described in Chapter 2 using the plasmid created by H. Bell[156]. The protein was concentrated to a concentration of 1.0 mM, confirmed by UV-Vis spectrophotometry, as shown in Figure 4.23.

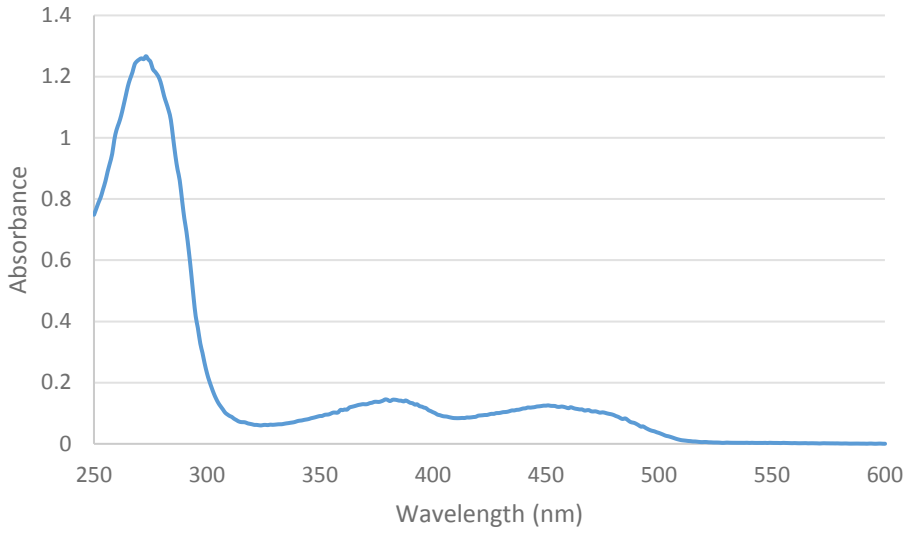

Figure 4.23 UV-visible spectrum of 1:100 dilution of H320F. Concentration calculated from absorbance of 0.125 at 450 nm with ε_{450} of 12,300 M^{-1} cm^{-1}.

The new PEG 8K conditions were used to set up hanging drop trays as described in Chapter 2. Two trays were set up, one co-crystallised with 1 mM L-Kyn, the other without L-Kyn. Crystals formed within 24 hours. Selected crystals were then soaked in 1 µL drops comprising of mother liquor stabilised with extra PEG 8K and 1 mM L-Kyn. After a further 24 hours individual crystals were flash cooled in liquid N_2 and sent to the Diamond light source.

Using beamline i03 data were collected. All crystals exhibited $P12_11$ spacegroup with dimensions similar to those seen in Table 3.4.

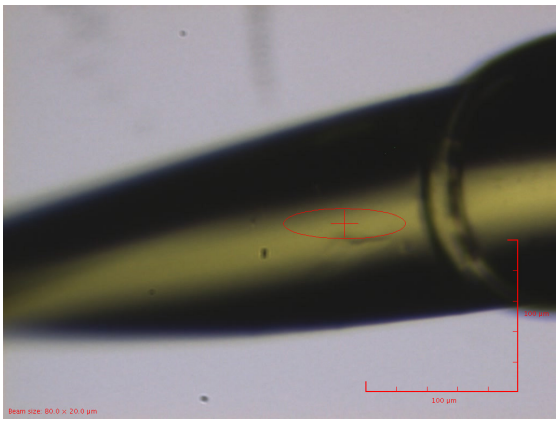

Figure 4.24 Mounted crystal of substrate-exposed H320F used for data collection.

From the crystal shown in Figure 4.24 2400 images were collected with beamline transmission 20 %, exposure of 0.1 s and oscillation of 0.15°. The statistics shown in Table 4.12 were calculated using the programs *iMosflm* and *Scala*[159, 160].

Resolution (Å)	R_{merge}	Mean (I)/σ(I)	Completeness (%)	Multiplicity
34.54-1.89	0.049	21.3	99.8	6.0

Table 4.12 Data collection statistics for a substrate-exposed H320F crystal. All calculated from overall data, for full data collection statistics see Appendix 7.1.4.

Refinement and model editing were performed as for dm2, described in section 2.8. The structure was solved with 2 molecules in the ASU and a solvent content of 47.2 %.

4.5.2. Discussion of substrate-bound H320F dm2 structure

Each molecule contained residues 6-457 and a total of 733 waters were modelled in the ASU. The final model was refined to give R and R_{free} values of 0.1550 and 0.2005, respectively, to a resolution of 1.89 Å. The full statistics can be seen in Appendix 7.1.4. The movement in the C-terminus was also exhibited between the two molecules within the ASU, as shown in Figure 4.25.

Figure 4.25 Crystal structure of substrate-bound H320F showing both molecules overlaid with one another. Molecule A is coloured in cyan, molecule B is coloured violet. RMSD = 1.255 Å from 437 aligned Cα atoms.

Molecule	Mean B-factor of C-terminus/Mean B-factor of structure
Closed	1.1
Open	1.9

Table 4.13 Average B-factor for residues 370-400 of substrate-bound H320F divided by average B-factor of overall structure.

The C-terminus shift is shown in Figure 4.25 is similar to that seen in dm2, as is the movement of residues 240-290. The B-factors shown in Table 4.13 also exhibit the same pattern of the closed molecule having the lower B-factors in the C-terminus.

The positioning of F320 is very similar to that of H320 in dm2 (the rings are superposable). The similarity between the two residues makes it difficult to ensure presence of a phenylalanine rather than a histidine. When using both 2Fo-Fc and Fo-Fc maps the results are not definitive, as shown in Figure 4.26.

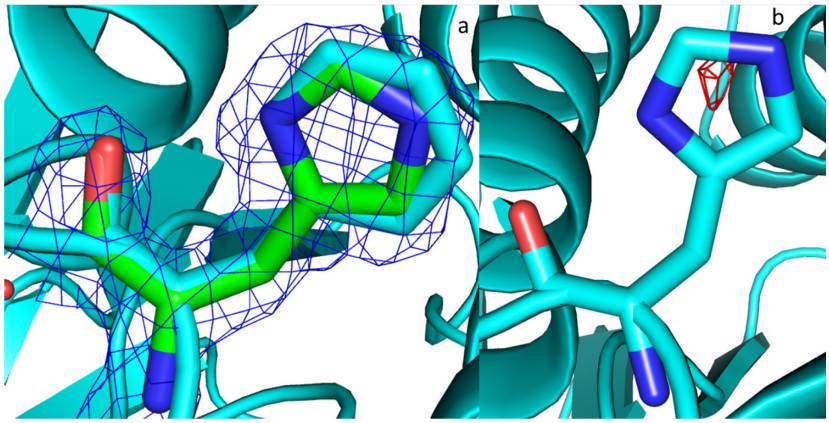

Figure 4.26 a. Crystal structure of substrate-bound H320F overlay with dm2 showing residues 320. 2Fo-Fc map is shown in blue (contoured at 0.50 eÅ^{-3}, 1.5 σ). H320F coloured cyan, dm2 coloured green. **b.** Crystal structure of substrate-bound H320F, modelled with His as residue 320. Fo-Fc map is shown in red (contoured at -0.28eÅ^{-3}, 2.7 σ).

Despite the presence of a phenylalanine codon being confirmed by Sanger sequencing for the H320F gene the protein was tested using mass spectrometry to support the presence of the correct residue.

H320F protein prepared in section 4.5 and dm2 protein prepared in section 3.2.2 were analysed using liquid chromatography-mass spectrometry. 15 µL of each protein was loaded onto an Aeris Widepore 3.6 µm C4 50 x 2.1 mm column connected to a Waters ZSpray Lockspray electrospray ionisation source, this was then analysed using a Synapt™ G2 time of flight mass spectrometer. The spectra are shown in Figure 4.27.

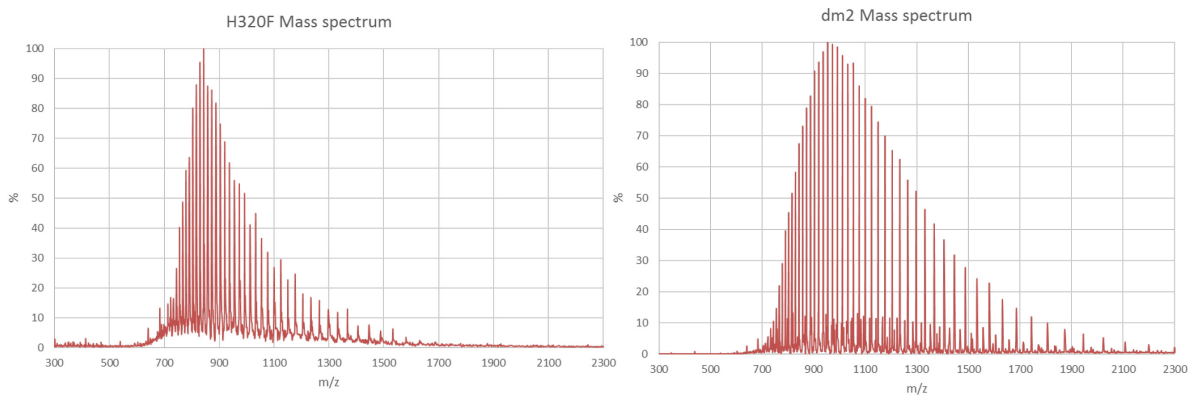

Figure 4.27 Spectra of H320F and dm2 using liquid chromatography electrospray ionisation mass spectrometry.

The mass of H320F is expected to be 10 Daltons greater than dm2. The mass of H320F was calculated to be 50564.3±1.5 Da while dm2 was calculated to be 50556.1±0.6 Da. This shows the appropriate mass difference between the two, confirming the substitution mutation of H320 to a phenylalanine.

The occupancy of L-Kyn is higher in H320F than in either R84K or Y404F, which matches their relative K_M values. The level of electron density surrounding L-Kyn is therefore greater, if not to quite the same extent as dm2, as shown in Figure 4.28. This is in line with H320F exhibiting a very similar binding affinity to dm2.

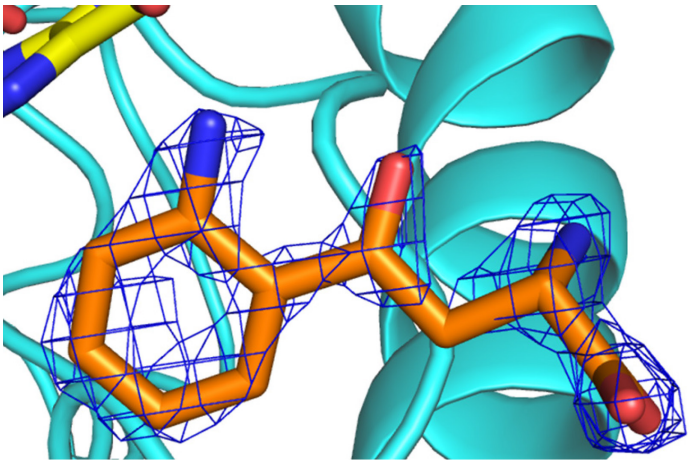

Figure 4.28 Crystal structure of crystal bound H320F showing L-Kyn and 2Fo-Fc map shown in blue (contoured at 0.43 eÅ-3, 1.51 σ).

When comparing the active sites of H320F to dm2 there is very little difference, as shown in Figure 4.29. The same interactions are seen between substrate and active site due to the almost identical positioning of the key residues and L-Kyn. As with dm2 the presence of L-Kyn is restricted to only one molecule, which is in the *closed* conformation.

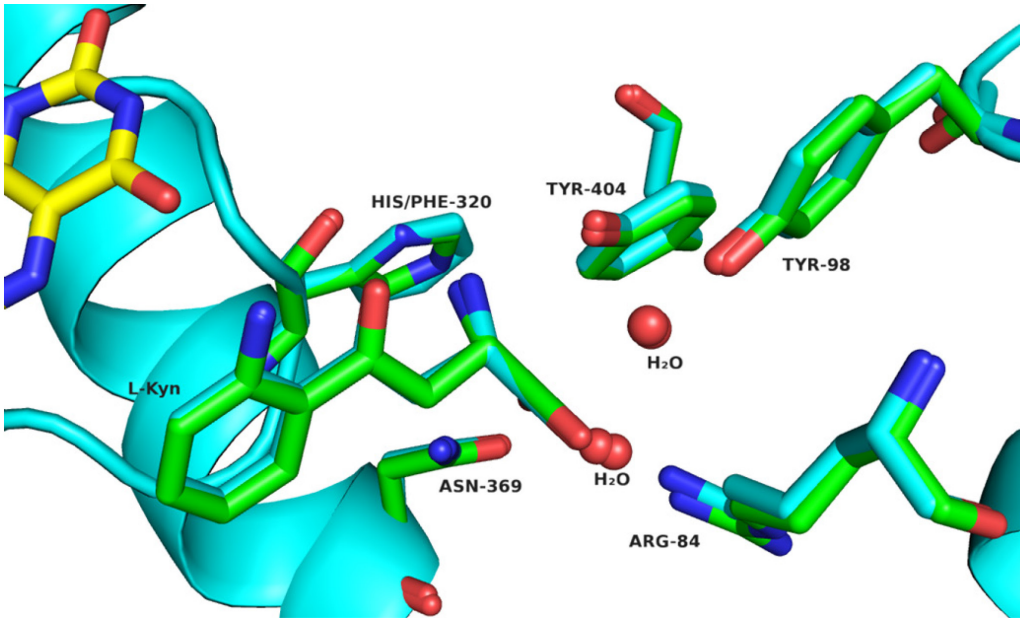

Figure 4.29 Crystal structure of substrate-bound H320F overlaid with substrate-bound dm2, showing L-Kyn and binding residues. H320F is coloured cyan, dm2 is coloured green. RMSD = 0.393 Å from 435 aligned Cα atoms.

There is very little change between H320F and dm2 in terms of H-bonding involving residue 320. This is because only the carbonyl of residue 320 interacts with the substrate. There are two water molecules that are within H-bonding distance from the nitrogen atoms of H320, but both of these form other interactions with surrounding groups including the carbonyl of residue 320, T408 and R403. Therefore we do not see a difference in the positioning of these waters in H230F.

4.6. Substrate-free H320F dm2

4.6.1. Crystal growth and data collection

Using the protein prepared in section 4.5 new crystal trays were prepared using the same method as described in section 4.5.1, with no exposure to L-Kyn. After 48 hours these substrate-free crystals were flash cooled in liquid N_2 and sent to Diamond Light Source. Using beamline i03 data were collected.

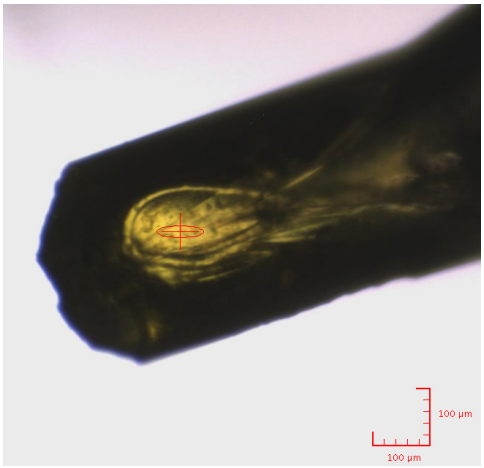

Figure 4.30 Mounted crystal of substrate-free H320F used for data collection.

From the crystal shown in Figure 4.30 1800 images were collected with beamline transmission 20 %, exposure of 0.15 s and oscillation of 0.15°. The statistics shown in Table 4.14 were calculated using the programs *iMosflm* and *Scala*[159, 160].

Resolution (Å)	R_{merge}	Mean (I)/σ(I)	Completeness (%)	Multiplicity
34.72-1.90	0.079	9.1	99.7	4.3

Table 4.14 Data collection statistics for a substrate-free H320F crystal. All calculated from overall data, for full data collection statistics see Appendix 7.1.4.

Refinement and model editing were performed as for dm2, described in section 2.8. The structure was solved with 2 molecules in the ASU and a solvent content of 47.6 %.

4.6.2. Discussion of substrate-free H320F dm2 structure

Each molecule contained residues 6-457 and a total of 465 waters were modelled in the ASU. The final model was refined to give R and R_{free} values of 0.1796 and 0.2272, respectively, to a resolution of 1.90 Å. The full statistics can be seen in Appendix 7.1.4. The movements in the C-terminus and residues 240-290 were also exhibited between the two molecules within the ASU, as shown in Figure 4.31.

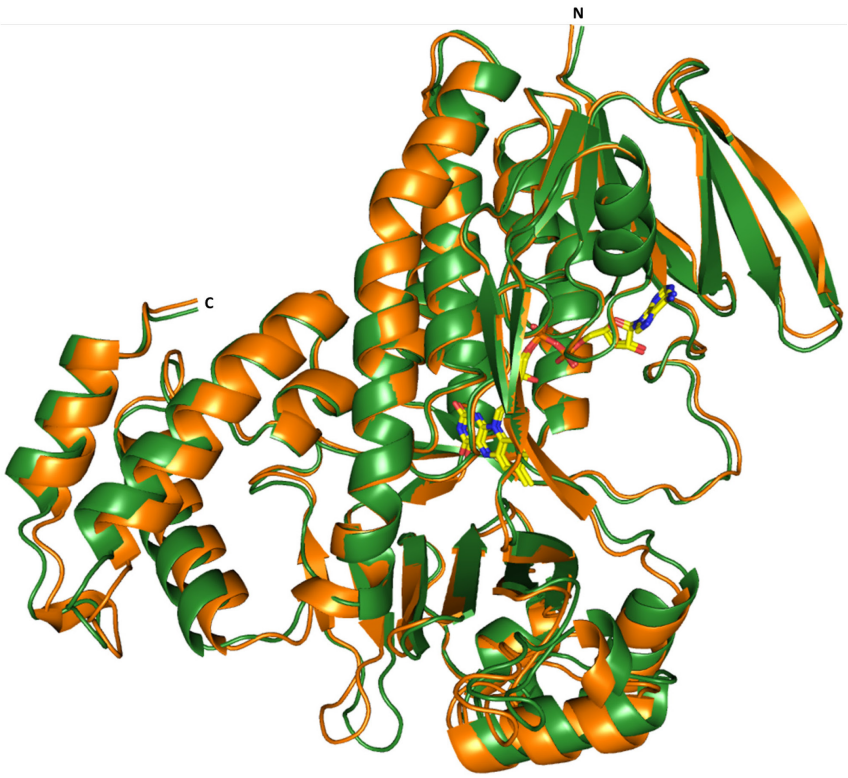

Figure 4.31 Crystal structure of substrate-free H320F showing both molecules overlaid with one another. Molecule A is coloured in orange, molecule B is coloured green. RMSD = 0.926 Å from 477 aligned Cα atoms.

Molecule	Mean B-factor of C-terminus/Mean B-factor of structure
Closed	1.7
Open	2.0

Table 4.15 Average B-factor for residues 370-400 of substrate-free H320F divided by average B-factor of overall structure.

The C-terminus shift observed in the substrate-free form of H320F is similar to that seen in dm2. The relation of B-factors between the two molecules are also as expected, as shown in Table 4.15, with the closed molecule having the lower value. The positioning of H320 is very similar between the two substrate-free molecules, as well as the substrate-bound model. This shows that the presence of substrate does not have an effect upon the positioning of this residue, which is consistent with it having no major role in substrate binding. An overlay of the four molecules is shown in Figure 4.32.

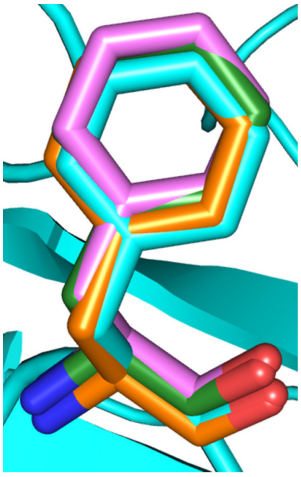

Figure 4.32 Crystal structure of H320F molecules, showing residue F320 from both molecules of substrate-bound and substrate-free structures. RMSD values were generated by aligning each molecule to the "closed" substrate-bound molecule (cyan). RMSD values of the overlays range from 0.351 – 1.255 Å, from at least 436 aligned Cα atoms in each case.

The active sites in the substrate-free and substrate-bound H320F molecules are also very similar, as shown in Figure 4.33.

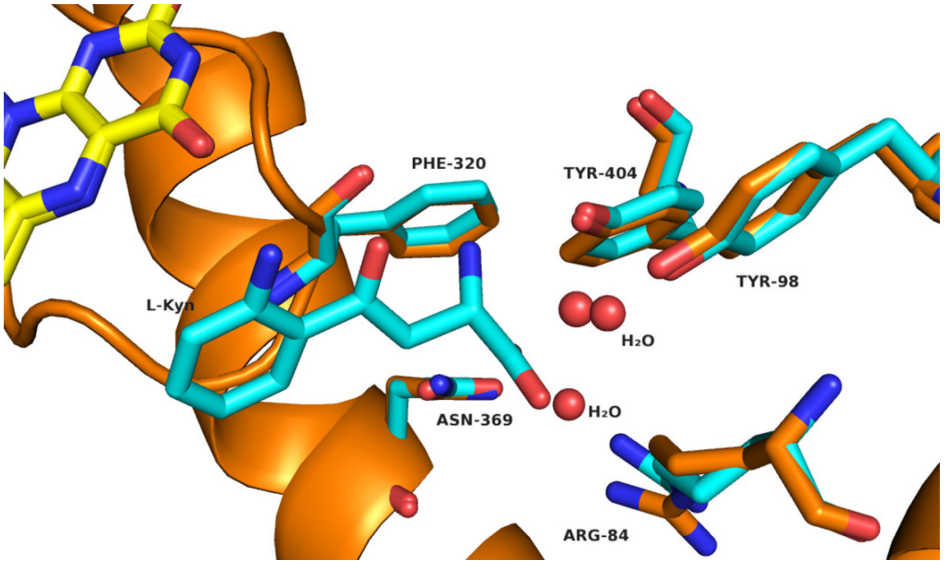

Figure 4.33 Crystal structure of substrate-free H320F (orange) overlaid with substrate-bound H320F (cyan), showing L-Kyn and binding residues. RMSD = 0.351 Å from 441 aligned Cα atoms.

H320F behaves very similarly to dm2 with very little difference seen in the active site regardless of the absence or presence of substrate.

As noted in this section there is great similarity in substrate binding, C-terminus movement and enzyme kinetics between dm2 and H320F. H320F has an analogous active site to hKMO and this similarity is encouraging in respect to dm2 acting as an accurate model for hKMO.

4.7. Conclusion

This chapter has examined the structural rationalisations for observations made in Michaelis-Menten kinetics in mutant forms of dm2. The influence of altering certain interactions within the active site vary between inconsequential to inactivating. The behaviour of substrate and residues within the active site provided reasoning for the Michaelis-Menten profile of R84K and Y404F, re-asserting the importance of these residues in substrate binding. The use of dm2 as a model for hKMO is also endorsed by the similarity demonstrated between dm2 and H320F.

The data surrounding the C-terminus are less straightforward. The structures within this chapter go against expectations; these being one *closed* conformation molecule with lower B-factors and one *open* molecule with higher B-factors. Firstly the molecules within the substrate free R84K structure are very similar, so similar that in some crystals they were considered symmetrical. This was also seen in some substrate free Y404F crystals. Secondly in substrate bound Y404F the B-factors of the *closed* molecule are shown to be slightly higher than the *open* molecule. This also contradicts the argument that the *closed* C-terminus is stabilised by L-Kyn.

From the data shown in substrate-free R84K it could be argued that the presence of substrate is possibly required to stabilise the C-terminus in the *closed* conformation. This is refuted by the presence of a *closed* molecule, with lower C-terminus B-factors in the structures of substrate free dm2, H320F and Y404F.

The movements of residues 240-290 are also atypical in this chapter. They behave as expected in most structures, even showing reduced movement to correlate with the reduced C-terminus movement in substrate-free R84K. This reduced movement, however, continues in structures where movement would be expected, in substrate-exposed and substrate-free Y404F. This makes a clear conclusion regarding the link between the two sections difficult to propose.

In general a pattern for the behaviour of the C-terminus is difficult to determine. Unlike in previous models the data presented in these latest chapters show there is not a clear relationship between L-Kyn presence and the stabilisation of one C-terminal conformation. The next chapter focuses upon the residues within the purported binding network between the mobile C-terminus and the rest of the protein.

Chapter 5: C-terminus residues

5.1. Introduction

In previous work by Martin Wilkinson of this group a network of interactions was proposed to be important for the stabilisation of the *closed* conformer of dm2[143]. This network was formed between residues in both the C-terminal domain and the remainder of the protein. These residues appeared to interact, closing a channel in the protein and excluding solvent from the active site to allow catalysis. It was also hypothesised that the presence of L-Kyn in the active site affected these residues causing this channel to close. In previous chapters it has been shown that the presence of L-Kyn has less effect upon the movement of the C-terminus than previously thought. This chapter investigates how the disruption of this potential network affects the movement of the C-terminus and the mechanism of dm2.

As seen in chapter 1, the residues identified as important to the putative network were R84, P97, E372, Y382 and R386. Substitution mutations were introduced, by Wilkinson, which targeted the residues R386 and E372, as these were seen to act as bridges between the two domains[143]. The effect of the substitutions R386K and E372T on the kinetic profile of the protein can be seen in Table 5.1.

Protein	K_M (μM)	k_{cat} (s^{-1})	k_{cat}/K_M (s^{-1} μM^{-1})
dm2	8.8 ± 1.5	8.9 ± 1.1	1.01
E372T dm2*	10.3 ± 0.7	8.7 ± 0.1	0.84
R386K dm2	17.0 ± 1.4	12.0 ± 0.9	0.71

Table 5.1 Summary of Michaelis-Menten kinetic data for mutant forms of dm2[143]. *E372T mutant also contained a L367R mutation.

There are no major differences between dm2, R386K and E372T in terms of Michaelis-Menten kinetics. This suggests that the residues in question are not directly involved in binding and that any disruption to the interactive network has little effect upon the efficiency of the enzyme.
In previous work the movement of the C-terminus in these mutants was also studied and it was found that the domain took up a *mixed* conformation, neither *open* nor *closed*[143]. These measurements were, however, hampered by the differing space groups of substrate-bound and substrate-free structures. Using the new PEG 8K crystallising conditions this issue is overcome. The previous two chapters have shown that the presence of L-Kyn is not necessary for dm2 to take up a *closed* conformation. This chapter investigates the presence of the putative inter-domain binding network and its effect upon the C-terminus.

This chapter will also present data of surface calculations for the *Pf*KMO variants analysed throughout this thesis. The measurement of the void volume or surface area of the active site of *Pf*KMO could

provide a more quantitative analysis of the role of the C-terminus when comparing the *open* and *closed* conformers.

5.2. Substrate-exposed R386K dm2

5.2.1. Crystal growth and data collection

R386K protein was expressed and purified as described in Chapter 2 using the plasmid created by Martin Wilkinson[143]. The protein was concentrated by centrifugation to a concentration of 430 μM confirmed by UV-Vis spectrophotometry, as shown in Figure 5.1.

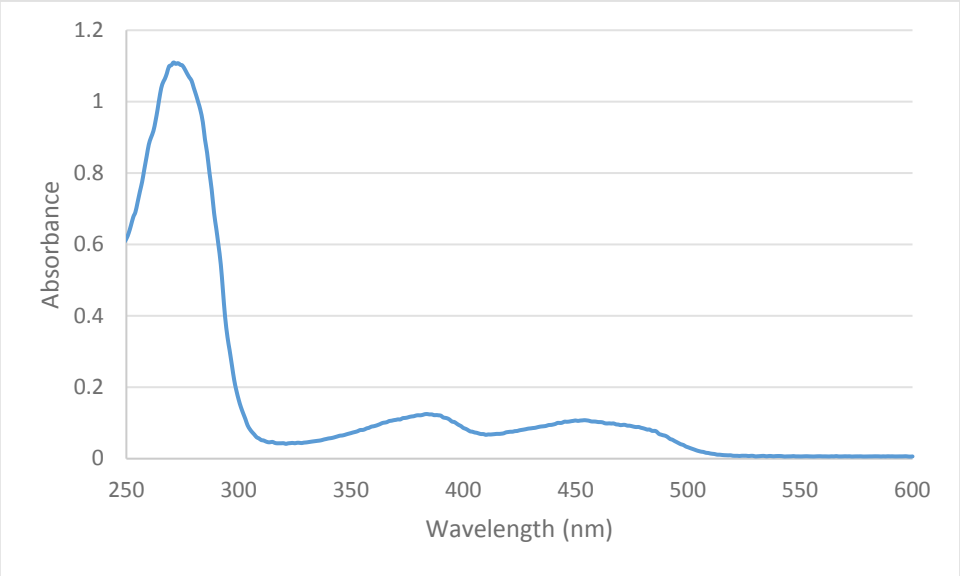

Figure 5.1 UV-Visible spectrum of 1:50 dilution of R386K. Concentration calculated from absorbance of 0.107 at 450 nm with ε_{450} of 12,300 M^{-1} cm^{-1}.

The new PEG 8K conditions were used to set up hanging drop trays as described in Chapter 2. Two trays were set up, one co-crystallised with 1 mM L-Kyn, the other without L-Kyn. Within 24 hours a number of large crystals had formed. Selected crystals were then soaked in 1 μL drops comprising of mother liquor stabilised with extra PEG 8K and 1mM L-Kyn. Crystals were then left for a further 24 hours and in some cases more than 72 hours. Individual crystals were flash cooled in liquid N_2 and sent to the Diamond light source. Data were collected using beamline i03. All crystals exhibited $P12_11$ spacegroup with similar parameters seen in Table 3.4.

Figure 5.2 Mounted crystal of substrate exposed R386K used for data collection.

From the crystal shown in Figure 5.2 1300 images were collected with beamline transmission 20.00 %, exposure of 0.1 s and oscillation of 0.15°. The statistics shown in Table 5.2 were calculated using the program *Xia2* and *Scala* [160, 161].

Resolution (Å)	R_{merge}	Mean (I)/σ(I)	Completeness (%)	Multiplicity
29.02-1.65	0.054	11.3	99.5	3.3

Table 5.2 Data collection statistics of a substrate-exposed R386K crystal. All calculated from overall data, for full data statistics see Appendix 7.1.5.

Refinement and model editing were performed as for dm2, described in section 2.8. The structure was solved with 2 molecules in the ASU and a solvent content of 47.7%. An overall image of the two molecules can be seen in Figure 4.3.

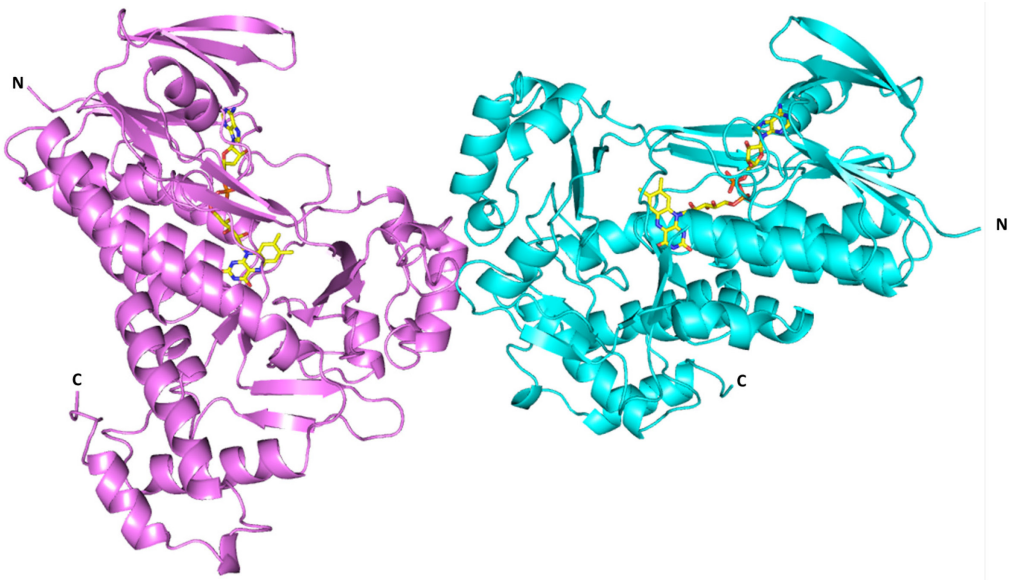

Figure 5.3 Crystal structure of substrate-exposed R386K showing both molecules in the ASU. Molecule A is coloured cyan, molecule B is coloured violet.

5.2.2. Discussion of substrate-exposed R386K dm2 structure

Each molecule contained residues 6-457 and a total of 572 waters were modelled in the ASU. The final model was refined to give R and R_{free} values of 0.1720 and 0.2095, respectively, to a resolution of 1.65 Å. The full statistics can be seen in Appendix 7.1.5. The movement in the C-terminus was also exhibited between the two molecules within the ASU, as shown in Figure 5.4.

Figure 5.4 Crystal structure of substrate exposed R386K, showing both molecules overlaid with one another. Molecule A is coloured cyan, molecule B is coloured violet. RMSD = 0.964 Å from 442 aligned Cα atoms.

Molecule	Mean B-factor of C-terminus/Mean B-factor of structure
Closed	1.4
Open	1.7

Table 5.3 Average B-factor for residues 370-400 of substrate exposed R386K divided by average B-factor of overall structure.

As seen in previous dm2 and dm2 mutant structures the movements of the C-terminus and residues 240-290 are present. The C-terminus B-factors, shown in Table 5.3, are also as expected with the *closed* molecule showing a lower value. The major difference between R386K and the enzymes of previous chapters is that L-Kyn is absent from the active site. This was found for all structures solved despite both co-crystallisation and crystal soaking being employed in addition to a much longer soaking time.

Fo-Fc maps of substrate-exposed and substrate-free structures were compared, much like R84K and Y404F. Ligand and water molecules were removed from the active site and the models were refined using *Refmac5*[164]. The resulting Fo-Fc maps can be seen in Figure 5.5.

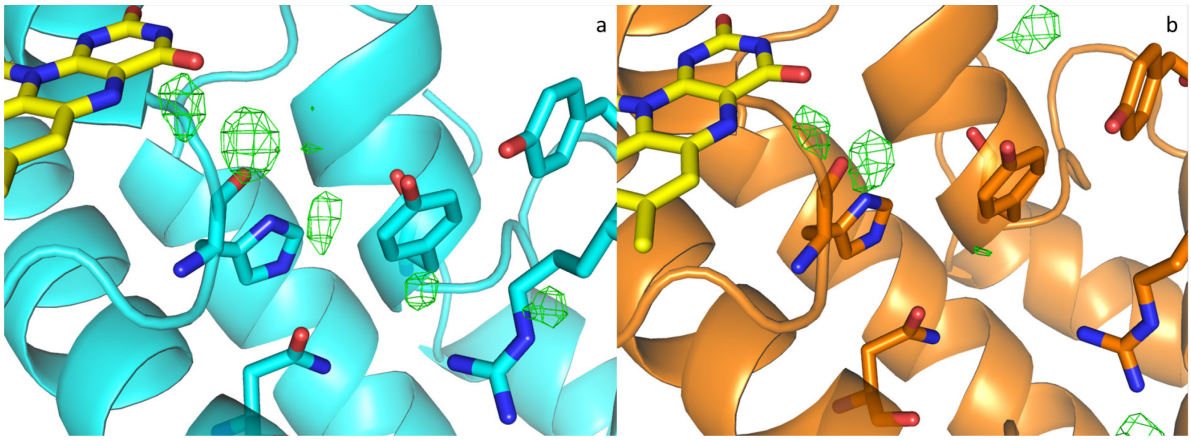

Figure 5.5 Crystal structures of R386K showing the active site. **a.** Substrate-exposed R386K with Fo-Fc map (contoured at 0.36 eÅ^{-3}, 2.97 σ) **b.** Substrate-free R386K with Fo-Fc map (contoured at 0.30 eÅ^{-3}, 2.96 σ).

Most of the green blobs exhibited in substrate-exposed R386K are also present in substrate-free R386K, including those in the expected position of L-Kyn. These peaks in the Fo-Fc maps model well as water molecules. These factors and the minimal differences between the two Fo-Fc maps suggest no L-Kyn occupancy in the substrate exposed structure.

Despite this the positioning of residues within the active site itself appears unaffected by the R386K mutation. A comparison with dm2 can be seen in Figure 5.6.

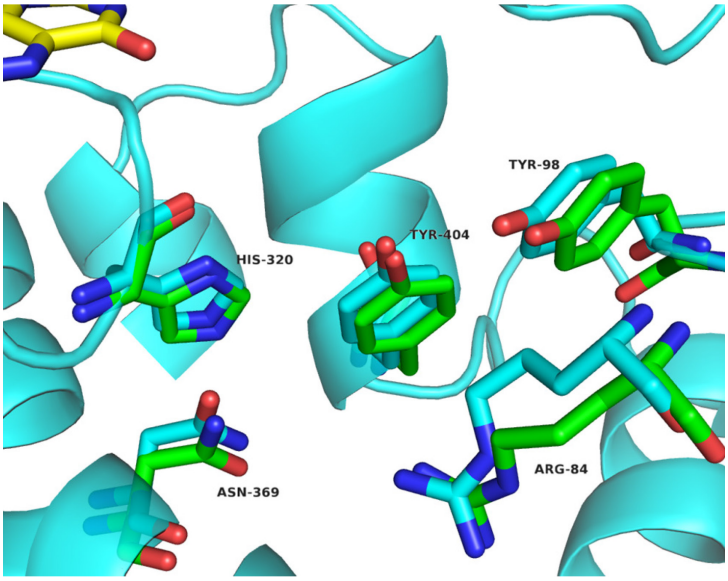

Figure 5.6 Crystal structures of substrate exposed R386K and substrate-bound dm2 overlaid, showing the active site. R386K is shown in cyan, dm2 is shown in green. RMSD = 1.537 Å from 437 aligned Cα atoms.

The lack of L-Kyn occupancy is difficult to explain, given R386K has a much lower K_M than Y404F for example. Even with the substitution of such an important residue evidence of L-Kyn occupancy is still observed in Y404F structures. This may suggest that the role of the C-terminus includes binding of substrate; though this is very difficult to argue with this evidence alone.

Due to the high B-factors within the surrounding domain it was necessary to increase the contouring of the 2Fo-Fc map to 1.0 σ in order to visualise the electron density of residue K386. This can be seen in Figure 5.7.

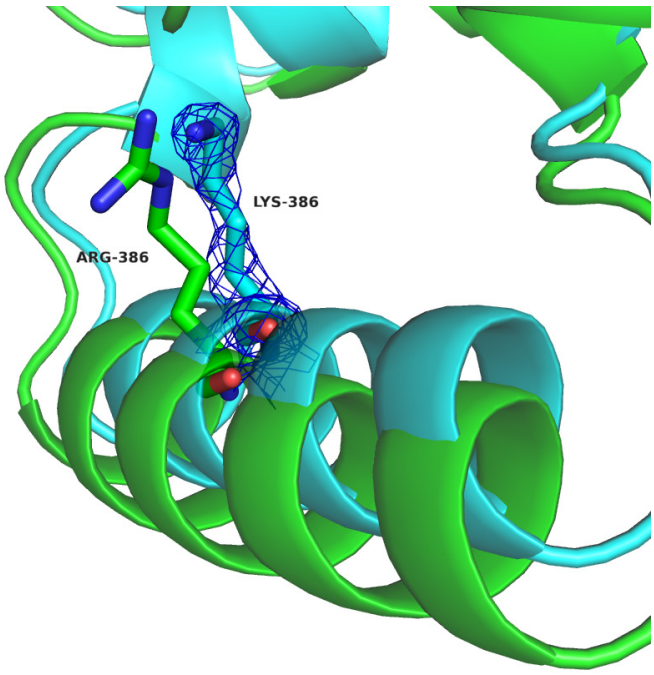

Figure 5.7 Crystal structures of substrate-exposed R386K and substrate-bound dm2 overlaid, showing residue 386. R386K is shown in cyan, dm2 is shown in green, 2Fo-Fc map is shown in blue (contoured at 0.36 eÅ^{-3}, 1.00 σ). RMSD = 1.537 Å from 437 aligned Cα atoms.

K386 is in a similar position to R386 of dm2, however the RMSD of the alignment is much higher than previous models. The distances between the two Cα atoms of residues 386 is 2 Å and the RMSD is 1.5 Å, making comparisons between the two models more difficult. Figure 5.4, however, shows a much smaller RMSD of 0.483 Å between the two R386K molecules, which are still distinguishable as *open* and *closed*. This suggests that the disruption of the putative network has no effect upon the ability of the protein to take up a *closed* conformation.

5.3. Substrate-free R386K dm2

5.3.1. Crystal growth and data collection

R386K crystals with no exposure to L-Kyn, prepared in section 5.2, were analysed at the same time as substrate-exposed crystals. Using beamline i03 data were collected.

Figure 5.8 Mounted crystal of substrate free R386K used for data collection.

From the crystal shown in Figure 5.8 1900 images were collected with beamline transmission 20.00 %, exposure of 0.1 s and oscillation of 0.10°. The statistics shown in Table 5.4 were calculated using the programs *iMosflm* and *Scala*[159, 160].

Resolution (Å)	R_{merge}	Mean (I)/σ(I)	Completeness (%)	Multiplicity
68.19-1.89	0.066	8.9	99.4	2.9

Table 5.4 Data collection statistics of a substrate-free R386K crystal. All calculated from overall data, for full data statistics see Appendix 7.1.5.

Refinement and model editing were performed as for dm2, described in section 2.8. The structure was solved with 2 molecules in the ASU and a solvent content of 48.5%.

5.3.2. Discussion of substrate-free R386K dm2 structure

Each molecule contained residues 6-457 and a total of 410 waters were modelled in the ASU. The final model was refined to give R and R_{free} values of 0.1770 and 0.2240, respectively, to a resolution of 1.89 Å. The full statistics can be seen in Appendix 7.1.5. The movements in the C-terminus and residues 240-290 were also exhibited between the two molecules within the ASU, as shown in Figure 5.9.

Figure 5.9 Crystal structure of substrate-free R386K, showing both molecules overlaid with one another. Molecule A is coloured green, molecule B is coloured orange. RMSD = 0.877 Å from 446 aligned Cα atoms.

Molecule	Mean B-factor of C-terminus/Mean B-factor of structure
Closed	1.6
Open	1.9

Table 5.5 Average B-factor for residues 370-400 of substrate-free R386K divided by average B-factor of overall structure.

The substrate-free R386K follows suit in exhibiting an *open* and *closed* conformer and having lower B-factors for the *closed* molecule, see Table 5.5. An overlay of residue 386 with dm2 can be seen in Figure 5.10.

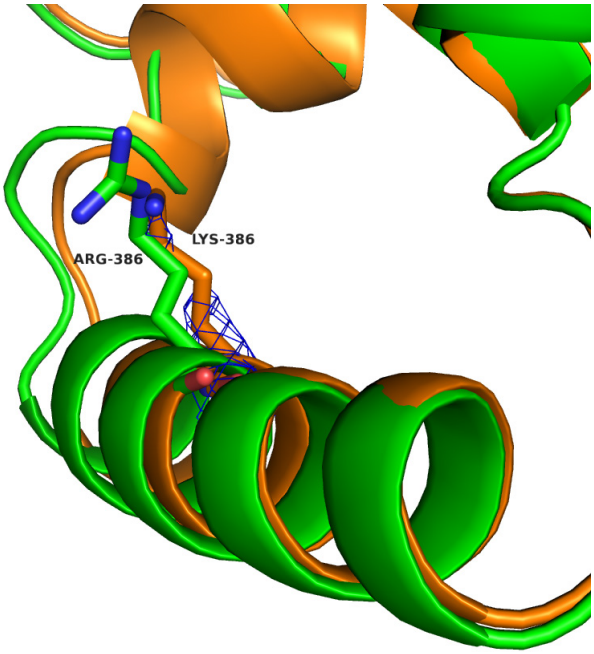

Figure 5.10 Crystal structures of substrate-free R386K and substrate-bound dm2 overlaid, showing residue 386. R386K is shown in orange, dm2 is shown in green, 2Fo-Fc map is shown in blue (contoured at 0.30 eÅ^{-3}, 1.00 σ). RMSD = 0.618 Å from 444 aligned Cα atoms.

In previous work it was postulated that the movement of the C-terminus of dm2 was controlled by the presence of L-Kyn, and this was enacted through a network of interactions. The data presented here appear to counter both of these hypotheses. Substrate-free R386K is capable of taking up the same *closed* conformation as the substrate-bound dm2 despite disruption of the putative interactive network, and having never been exposed to L-Kyn.

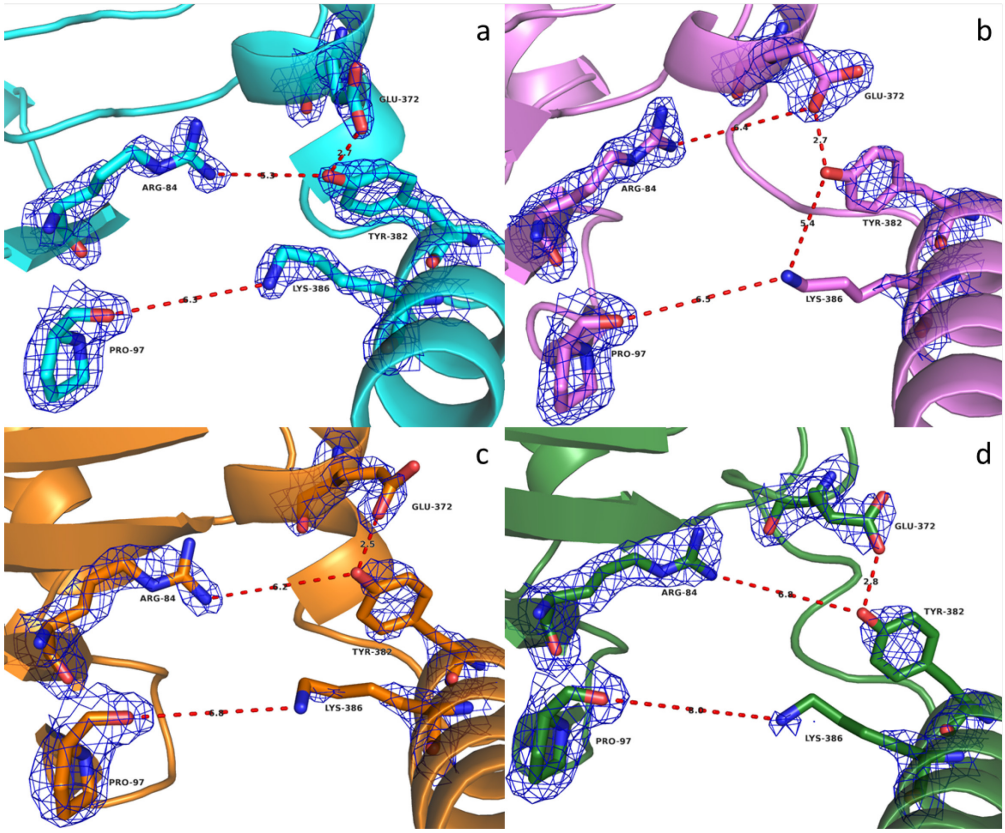

Figure 5.11 Crystal structure showing residues involved in the putative interacting network of the C-terminus of R386K **a.** Closed molecule of L-Kyn-exposed R386K **b.** Open molecule of L-Kyn-exposed R386K **c.** Closed molecule of L-Kyn-free R386K **d.** Open molecule of L-Kyn-free R386K. 2Fo-Fc maps are shown in blue (L-Kyn exposed molecules contoured at 0.36 eÅ^{-3}, 1.0 σ, L-Kyn free molecules contoured at 0.30 eÅ^{-3}, 1.0 σ). Distances between atoms are in Å and shown as red dashed lines.

Figure 5.11 shows the positioning of the other residues within the putative interactive network in all the R386K molecules. The overall positioning of K386 is similar to R386 seen in dm2. However, much like its counterpart in dm2 the residue is very mobile. The lysine appears to be a comparable distance from P97 as R386 is in dm2 in the *closed* molecule. However, given the mobility and relatively inconsistent positioning of R386 in dm2 it is unsurprising K386 follows a similar trend. The mobility of the residue therefore leads to very sparse density shown in the 2Fo-Fc map. Consequently comparisons between the two residues are difficult to make. The conclusions we can form are twofold. Firstly the distances measured between the residues are similar to dm2 in that they are too large to form any meaningful interactions. Secondly the appearance of *closed* or *open* conformations seems unaffected by this mutation, as is the kinetic ability of the protein.

5.4. Substrate-exposed E372T dm2

5.4.1. Crystal growth and data collection

The extra L367R mutation present in the kinetic testing of E372T in Table 5.1 was corrected for crystallography. E327T protein was expressed and purified as described in Chapter 2 using the plasmid created by Martin Wilkinson[143]. The protein was concentrated by centrifugation to a concentration of 585 μM confirmed by UV-Vis spectrophotometry, as shown in Figure 5.12.

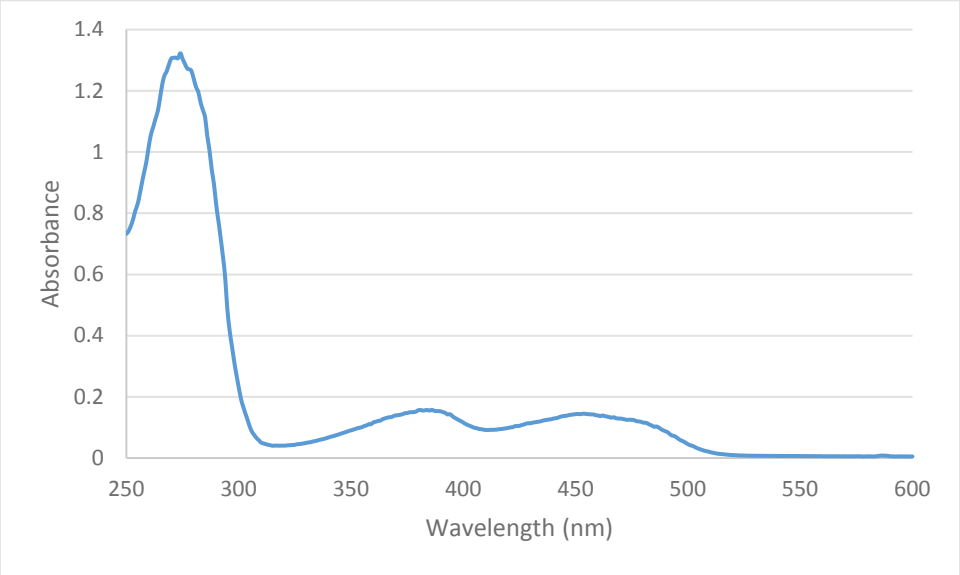

Figure 5.12 UV-Visible spectrum of 1:50 dilution of E372T. Concentration calculated from absorbance of 0.144 at 450 nm with ε_{450} of 12,300 M^{-1} cm^{-1}.

The new PEG 8K conditions were used to set up hanging drop trays as described in Chapter 2. Two trays were set up, one co-crystallised with 1 mM L-Kyn, the other without L-Kyn. After 24 hours only one crystal had formed from both the trays. Two more trays were prepared, one with 1 mM L-Kyn, one without L-Kyn. The crystal was crushed and used to seed the drops of the new trays, using the method described in Chapter 2. Crystals formed after 24 hours. Several dilutions of the seeding solution were used with the 1:1000 dilution yielding the largest and highest diffracting crystals. Selected crystals were then soaked in 1 μL drops comprising of mother liquor stabilised with extra PEG 8K and 1 mM L-Kyn. Crystals were then left for a further 24 hours and in some cases more than 72 hours. Individual crystals were flash cooled in liquid N_2 and sent to the Diamond light source. Data were collected using beamline i02.

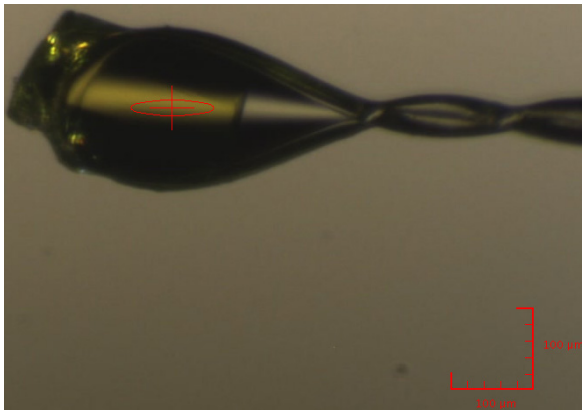

Figure 5.13 Mounted crystal of substrate-exposed E372T used for data collection.

From the crystal shown in Figure 5.13 3600 images were collected with beamline transmission 20.00 %, exposure of 0.1 s and oscillation of 0.05°. The statistics shown in Table 5.6 were calculated using the programs *iMosflm* and *Scala*[159, 160].

Resolution (Å)	R_{merge}	Mean (I)/σ(I)	Completeness (%)	Multiplicity
67.69-1.79	0.047	11.0	99.5	3.1

Table 5.6 Data collection statistics of a substrate-exposed E372T crystal. All calculated from overall data, for full data statistics see Appendix 7.1.6.

Refinement and model editing were performed as for dm2, described in section 2.8. The structure was solved with 2 molecules in the ASU and a solvent content of 47.4%.

5.4.2. Discussion of substrate-exposed E372T dm2 structure

The model contained residues 6-459 in molecule A, 6-457 in molecule B and 356 waters in the ASU. The final model was refined to give R and R_{free} values of 0.1859 and 0.2297, respectively, to a resolution of 1.79 Å. The full statistics can be seen in Appendix 7.1.6. The movement of the C-terminus was exhibited between the two molecules within the ASU, as shown in Figure 5.14. Table 5.7 describes the average B-factors for the C-termini. The movement of residues 240-290 is also seen but is significantly decreased, unlike the shift in the C-terminus. This supports the idea that the movement is not caused by the C-terminus but is simply a product of motility and thermal motion.

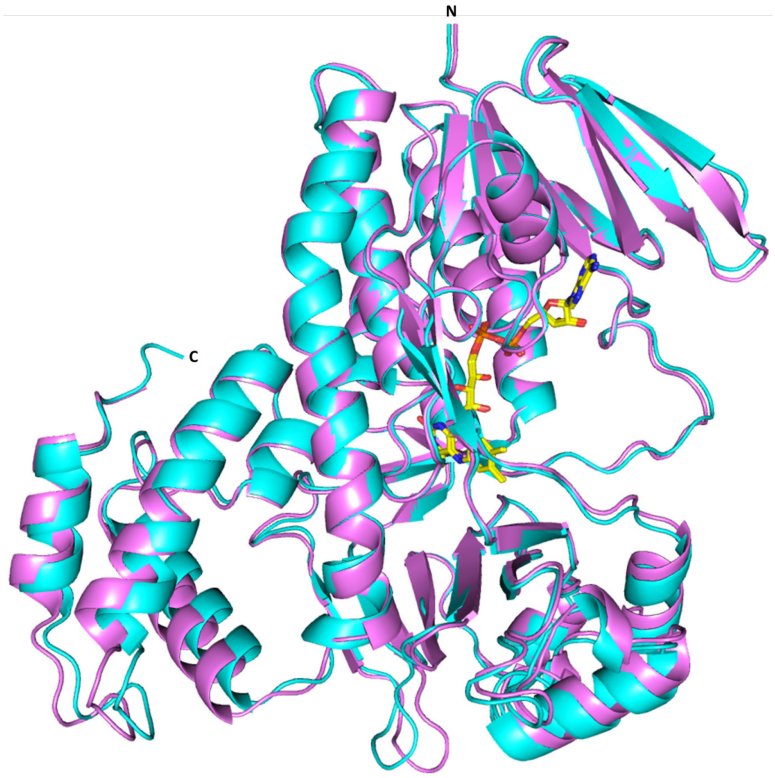

Figure 5.14 Crystal structure of substrate-exposed E372T, showing both molecules overlaid with one another. Molecule A is coloured cyan, molecule B is coloured violet. RMSD = 0.827 Å from 448 aligned Cα atoms.

Molecule	Mean B-factor of C-terminus/Mean B-factor of structure
Closed	1.6
Open	1.9

Table 5.7 Average B-factor for residues 370-400 of substrate-exposed E372T divided by average B-factor of overall structure.

Both the movement of the C-terminus and the lower B-factor in the *closed* molecule are present within the model. As with R386K L-Kyn is undetected in the active site, despite extended soaking time. Comparisons of the Fo-Fc maps of substrate-exposed and substrate-free E372T structures are shown in Figure 5.15. Once again these exhibit very little difference showing no evidence of L-Kyn occupancy.

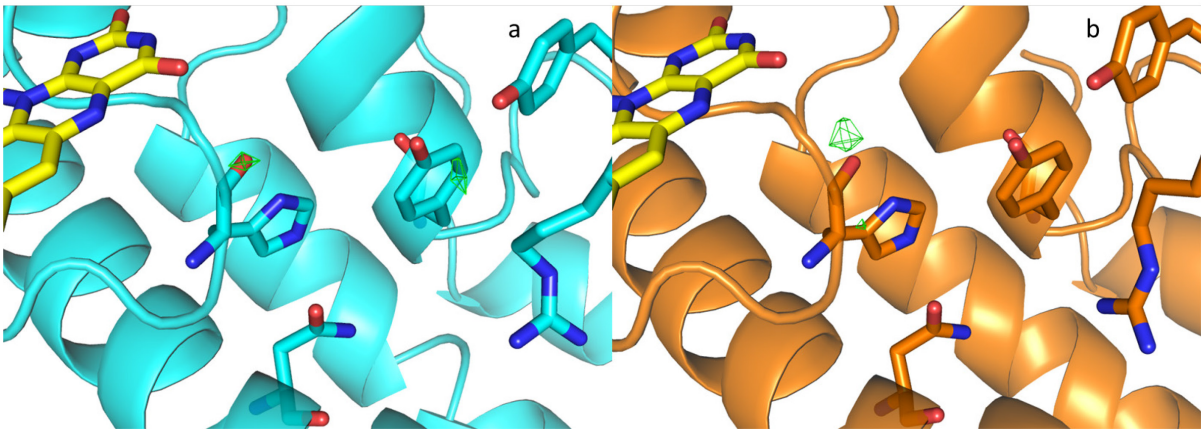

Figure 5.15 Crystal structures of E372T showing the active site. **a.** Substrate-exposed E372T with Fo-Fc map (contoured at 0.27 eÅ^{-3}, 2.99 σ) **b.** Substrate-free E372T with Fo-Fc map (contoured at 0.28 eÅ^{-3}, 3.00 σ).

As with R386K the positioning of active site residues is very similar to dm2, as shown in Figure 5.16.

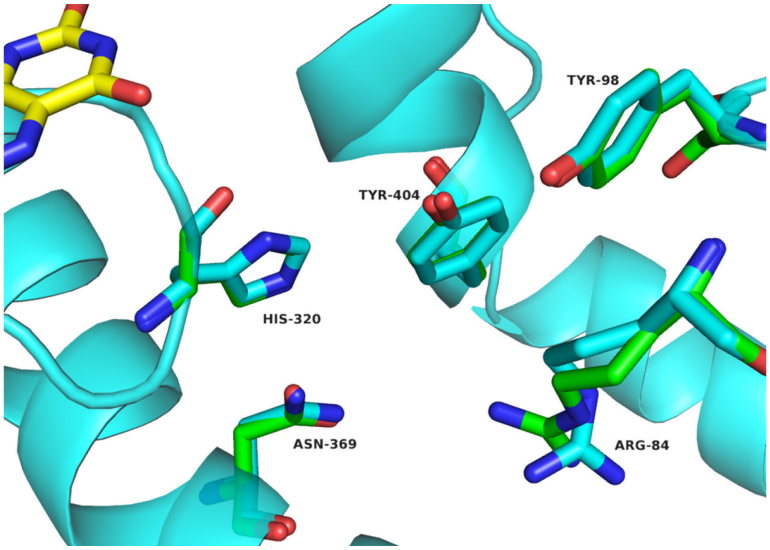

Figure 5.16 Crystal structures of substrate-exposed E372T and substrate-bound dm2 overlaid, showing the active site. E372T is shown in cyan, dm2 is shown in green. RMSD = 0.582 Å from 435 aligned Cα atoms.

The B-factors of the molecule overall are very high and the difficulty in crystallising the protein may indicate a less static, more dynamic protein, particularly in the C-terminus. This is difficult to quantify, however, since comparing absolute B-factors between different models of various resolutions can be problematic.

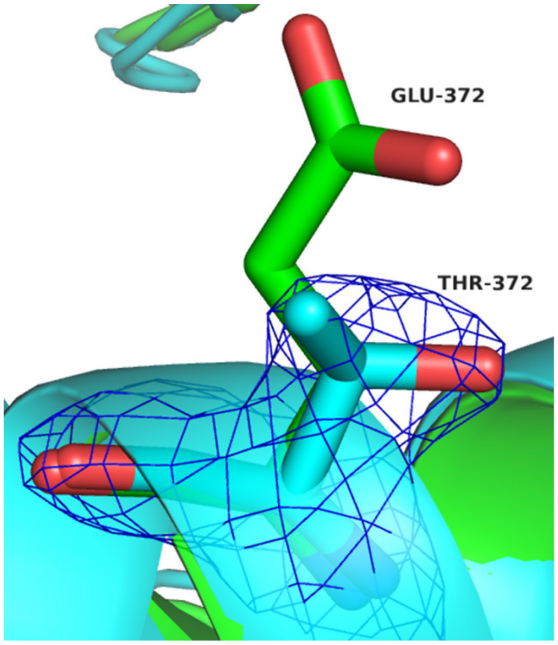

Figure 5.17 Crystal structures of substrate-exposed E372T and substrate-bound dm2 overlaid, showing residue 372. E372T is shown in cyan, dm2 is shown in green, 2Fo-Fc map is shown in blue (contoured at 0.26 eÅ^{-3}, 1.00 σ). RMSD = 0.582 Å from 435 aligned Cα atoms.

As shown in Figure 5.17, the positioning of the threonine and direction of the alcohol group are very different to the acid group of the glutamate. This means any potential interactivity E372 might have been involved with would be greatly diminished if not impossible with the threonine side chain. As shown in Figure 5.14 the molecule is still capable of forming a *closed* conformation.

5.5. Substrate-free E372T dm2

5.5.1. Crystal growth and data collection

E372T crystals with no exposure to L-Kyn, prepared in section 5.4, were analysed at the same time as substrate-bound crystals. Using beamline i02 data were collected.

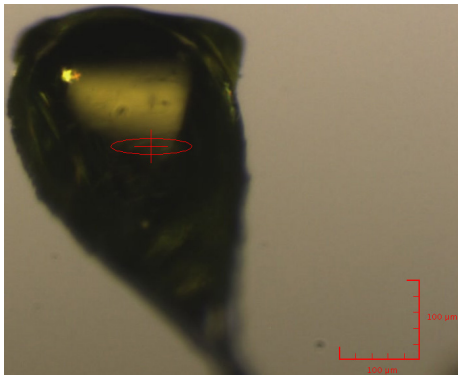

Figure 5.18 Mounted crystal of substrate-free E372T used for data collection.

From the crystal shown in Figure 5.18 2000 images were collected with beamline transmission 20.00 %, exposure of 0.05 s and oscillation of 0.10°. The statistics shown in Table 5.8 were calculated using the programs *iMosflm* and *Scala*[159, 160].

Resolution (Å)	R_{merge}	Mean (I)/σ(I)	Completeness (%)	Multiplicity
67.30-1.88	0.085	6.3	99.1	3.4

Table 5.8 Data collection statistics of a substrate-free E372T crystal. All calculated from overall data, for full data statistics see Appendix 7.1.6.

Refinement and model editing were performed as for dm2, described in section 2.8. The structure was solved with 2 molecules in the ASU and a solvent content of 46.5%.

5.5.2. Discussion of substrate-free E372T dm2 structure

Each molecule contained residues 6-457 and a total of 314 waters were modelled in the ASU. The final model was refined to give R and R_{free} values of 0.1973 and 0.2435, respectively, to a resolution of 1.88 Å. The full statistics can be seen in Appendix 7.1.6. The movement in the C-terminus was also exhibited between the two molecules within the ASU, as shown in Figure 5.19. The movement of residues 240-290 is also observed but, as seen in section 5.4.2, is significantly less than expected when compared to the movement of the C-terminus.

Figure 5.19 Crystal structure of substrate-free E372T, showing both molecules overlaid with one another. Molecule A is coloured green, molecule B is coloured orange. RMSD = 0.888 Å from 452 aligned Cα atoms.

Molecule	Mean B-factor of C-terminus/Mean B-factor of structure
Closed	1.6
Open	1.8

Table 5.9 Average B-factor for residues 370-400 of substrate-free E372T divided by average B-factor of overall structure.

As shown in Figure 5.19 and Table 5.9 the model exhibits the C-terminus movement and lower B-factors in the *closed* conformation. In agreement with findings of the R386K molecule the disruption of any potential interactive network between the C-terminus and the rest of the protein has no effect on the protein's ability to achieve the *closed* conformation. The absence of L-Kyn also has no effect.

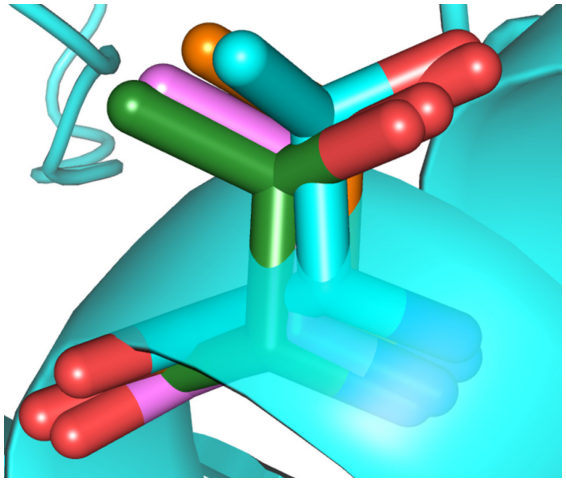

Figure 5.20 Crystal structure of E372T molecules, showing residue T372 from both molecules of substrate-exposed and substrate-free structures. RMSD values were generated by aligning each molecule to the “closed” substrate-bound molecule (cyan). RMSD values of the overlays range from 0.235 – 0.833 Å, from at least 444 aligned Cα atoms in each case.

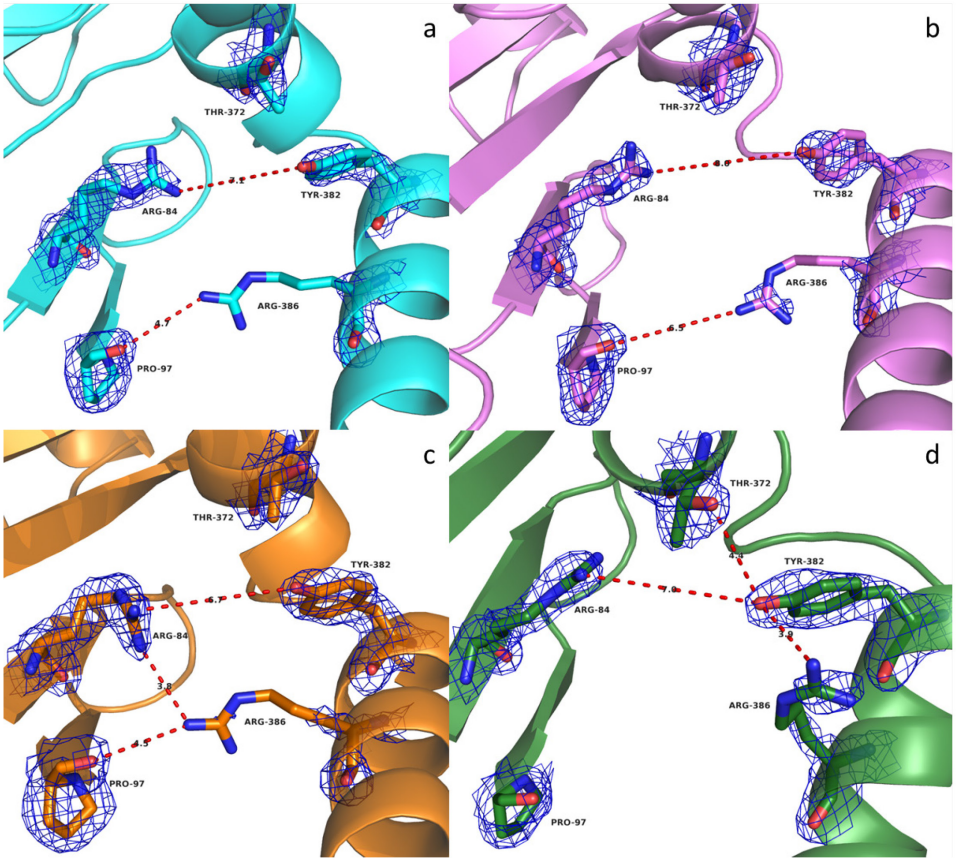

Figure 5.21 Crystal structure showing residues involved in the putative interacting network of the C-terminus of E372T **a.** Closed molecule of L-Kyn exposed E372T **b.** Open molecule of L-Kyn exposed E372T **c.** Closed molecule of L-Kyn free E372T **d.** Open molecule of L-Kyn free E372T. 2Fo-Fc maps are shown in blue (L-Kyn exposed molecules contoured at 0.26 eÅ^{-3}, 1.0 σ, L-Kyn free molecules contoured at 0.25 eÅ^{-3}, 1.0 σ). Distances between atoms are in Å and shown as red dashed lines.

As shown in Figure 5.20 and Figure 5.21 residue T372 is very similarly positioned in each molecule and plays no part in the putative interactive network. All distances depicted in Figure 5.21 are too large to indicate any meaningful interaction. This does not seem to affect the ability of the molecule to form a *closed* conformer.

The interactivity of the proposed network has not been significant in any structure throughout this work.

5.6. Surface calculations

The analysis of void volume and surface area within the active site of dm2 yielded inconsistent results, as seen in chapter 3. The *open* molecule of dm2 was expected to have a larger void volume and a larger surface area in the active site and at the opening of the channel. In the substrate-bound model the volume was very similar in each molecule and the surface area larger in the *open* molecule. In the substrate-free model both the volume and the surface area were larger in the *closed* molecule. The data were not sufficient to draw any conclusions. Therefore, surface calculations for all the mutants presented so far in this thesis were performed to compare their active sites. Table 5.10 presents the surface area and void volume of the active sites of the *open* and *closed* molecules of each protein.

Substrate-Bound molecule	Void Volume (Å^3)	Surface Area (Å^2)	Substrate-Free molecule	Void Volume (Å^3)	Surface Area (Å^2)
dm2					
Closed	376	413	*Closed*	383	449
Open	374	431	*Open*	332	382
R84K					
Closed	461	501	*Closed*	434	477
Open	447	509	*Open*	433	476
Y404F					
Closed	364	408	*Closed*	408	448
Open	438	487	*Open*	407	465
H320F					
Closed	335	393	*Closed*	370	424
Open	365	431	*Open*	359	422
R386K					
Closed	370	428	*Closed*	398	450
Open	398	444	*Open*	379	438
E372T					
Closed	398	444	*Closed*	385	445
Open	370	428	*Open*	335	404

Table 5.10 Void volume and surface area of the active site of both *open* and *closed* conformers of *Pf*KMO variants.

The data presented in Table 5.10 shows very little pattern or order. This matches most of the findings concerning the C-terminus of *Pf*KMO. In some cases the active site of the *closed* molecule has significantly smaller volume and surface area, such as substrate-bound H320F. Other examples show the *open* molecule with the smaller values, such as substrate-free dm2. Some have very similar values for both volume and surface area; some differ in just one criterion.

The difficulty in quantifying the surface area and volume of the active sites in these molecules may stem from the changes seen in the opening. As the C-terminus moves the extent of the opening differs. At times it is closed enough to act as a surface within the active site; other times the opening is large enough to become void volume and quickly becomes the surface of the protein. A very slight movement can result in large differences in the value of volume and surface area within the calculation.

In previous work the hypothesis of substrate induced fit required the presence of a clearly defined channel opening to allow substrate access and exclude solvent. Throughout this thesis the opening,

much like the movement of the C-terminus, has been less binary than expected. It may therefore be thought of as more of a cleft than an established channel. This makes measuring the extent of opening and closing very difficult. Given the nature of the C-terminus movement, however, this may not be necessary. The study of the C-terminus throughout this thesis has shown no evidence toward substrate-induced fit or the need for an interactive network between the C-terminus and the rest of the protein. These findings still hold true despite the difficulty in surface area and void volume calculations.

Chapter 6: Conclusions

The most important result of this work is the first high-resolution structure of substrate-bound KMO, determined at a resolution of 1.5 Å. The interactions between L-Kyn and the active site of dm2 can form a template for future inhibitor design. This has already impacted the design of potential inhibitors as noted in the soon to be published paper by Hutchinson *et al*.[133, 154].

Residues of the active site have also been successfully investigated including R84, Y404 and H320. The kinetic behaviour of these dm2 variants has been explained through structural findings. The importance of residues R84 and Y404 have been ratified and the introduction of the H320F substitution has afforded the structure of an active site homologous to that of hKMO. The similarity in binding observed in dm2 and H320F also affirms *Pf*KMO as a suitable model for the study of hKMO and as a basis for inhibitor design. Despite exhaustive efforts the occupancy of substrate has been low in R84K and Y404F structures and required altered contouring of the 2Fo-Fc maps. Given the relatively high K_M of L-Kyn followed by further alteration of the active site this, perhaps, is not surprising. Supporting factors for the case of low substrate occupancy include the fact that the substrate can be modeled with 100% occupancy, the differences shown between the Fo-Fc maps of substrate-exposed and substrate-free models and the supporting polder maps.

The findings concerning the movement of the C-terminus are not straightforward. The most clear-cut conclusion surrounding the C-terminus movement is disproving the hypothesis that a ligand is required for dm2 to achieve a *closed* form. The work also disbands the proposed reliance upon an interactive network between the C-terminus and the rest of the protein in order to take up the *closed* conformation. Furthermore there is no evidence to imply the positioning of the C-terminus is caused by crystal contacts within the unit cell.

The large body of data recorded throughout this work supports a proposal that the movement of the C-terminus is not controlled but is simply the result of high mobility and thermal motion. The positioning of the C-terminus follows no pattern throughout; it differs from crystal to crystal and has been effectively random. In many of the proteins studied certain datasets reduced the size of the ASU to include only one molecule. This was due to the molecules being similar enough to do so, thus implying the difference between C-termini was minimal. Other datasets of the same molecule have shown significant distinction between the two.

This lack of clarity is echoed in the surface calculations for the active site and opening cleft. Given the proximity of the surface of the protein small differences in C-terminus positioning resulted in large differences in surface area and volume values. The calculations yielded unpredictable results that follow no clear pattern. These results, however, do not take away from the main conclusions concerning the C-terminus this thesis has put forward.

The movement of residues 240-290 is also haphazard enough to make any conclusions difficult. Evidence exists to support a connection between the motion of the C-terminus and these residues; given their movements are matched in a number of structures. This is, however, countered in other structures where movement of residues 240-290 is significantly less than expected. This movement may be the result of high mobility and thermal motion, and we are capturing snapshots of this motion. The motion may also be connected to movement of the C-terminus. The data presented does not allow for a strong conclusion either way. Given the positioning of these residues compared to the active site and the opening caused by the C-terminus movement the connection may not be particularly important anyway.

The level of substrate occupancy also appears to be linked to the movement of the C-terminus. Multiple iterations of dm2 have been presented in this thesis; the molecules in the *closed* form often show evidence of substrate binding. The molecules in the *open* form show much less evidence of substrate binding. This can be explained as when dm2 is in the *closed* state the substrate is trapped within the molecule, when the molecule is *open* the L-Kyn can diffuse out of the active site, thus reducing occupancy. This could be explained as conformational selection; wherein the fluid dynamism of KMO leads to the *closed* state and more effective binding of substrate. The findings in this thesis are in keeping with the findings of Changeux *et al.* in that the majority of data supports conformational selection over induced fit[176]. In summary the movement of the C-terminus may yet relate to substrate binding and that the relationship remains difficult to fully define.

The work of industrial collaborators has also shown the presence of an *open* and *closed* molecule within the ASU. Using the same crystallising conditions both apo and inhibitor-bound structures showed the same movement between the two chains. One example is the competitive ligand GSK428. The IC_{50} of GSK428 is 60 nM implying a K_i two orders of magnitude lower than the K_M of L-Kyn[133, 154]. This resulted in ligand binding in both molecules. Despite the low K_i the ligand still showed slightly different interactions within the active site. The more *closed* chain B shows more indirect water-bridged interactions, reminiscent of L-Kyn, than the *open* chain A. This matches with the findings presented in this work and could implicate the C-terminus in ligand binding.

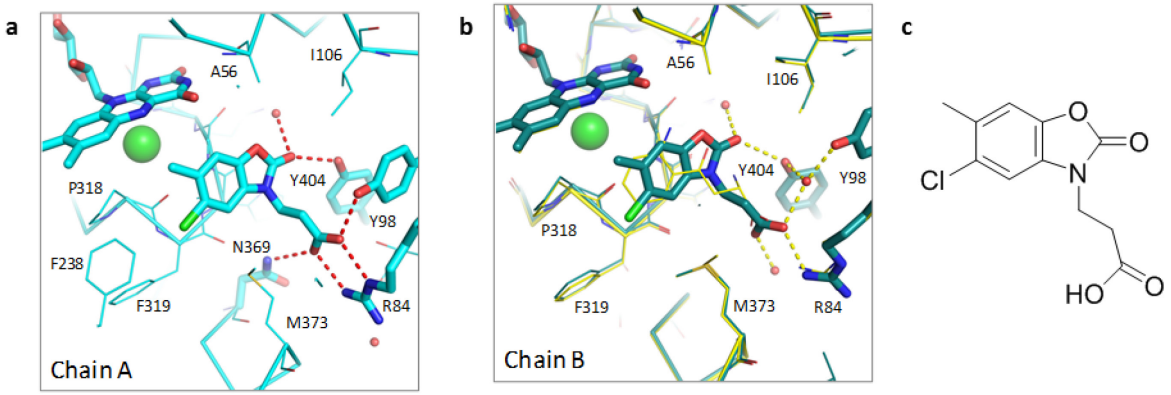

Figure 6.1 Crystal structure of dm2 bound with inhibitor GSK428. **a.** Binding site of chain A of the ASU shown in cyan **b.** Binding site of chain B of the ASU shown in blue. **c.** Structure of GSK428. Interactions are shown as dashed lines. The green sphere represents a chloride ion. Image adapted from the paper by Hutchinson et al. with permission from the authors[133]. (PDB: 5NAB)

The failure to capture a substrate molecule within R386K and E372T can also be used to argue the link between the C-terminus and the binding of L-Kyn. The K_M of R386K and E372T are so similar to dm2 that one could expect the substrate to bind more easily, as seen in dm2 or H320F. The higher mobility of the C-terminus in R386K and E372T, implied by their higher B-factors, may mean L-Kyn is more likely to diffuse out of the active site.

It is difficult to examine this theory with the data we currently have. In order to investigate the role of C-terminus movement in substrate binding it would be necessary to reduce its mobility. This could involve fixing the C-terminus of dm2 in an *open* or *closed* state, preferably both. Structural and kinetic study on this subject could improve our understanding of the role of the C-terminus. Introduction of cysteine substitutions could cause di-sulfide bonds to form between the C-terminus and the other domain of the protein, closing the channel to the active site. This may also have the effect of decreasing the solubility of the protein, due to unwanted di-sulfide bonds between molecules. This would lead to aggregation as seen in the wild type *Pf*KMO. Given that dm2 already has two cysteine residues replaced, additional cysteines could prove problematic.

The further work of industrial collaborators also led to a new kind of KMO inhibitor that may point towards a flavin movement within KMO and type A flavoprotein aromatic hydroxylase proteins in general. By adding an alkoxyl-linked pyridyl group the affinity of inhibitor GSK891, shown in Figure 6.2, increased dramatically. It was established from the crystal structure the pyridine π-stacks with the flavin group causing a tilt in the FAD.

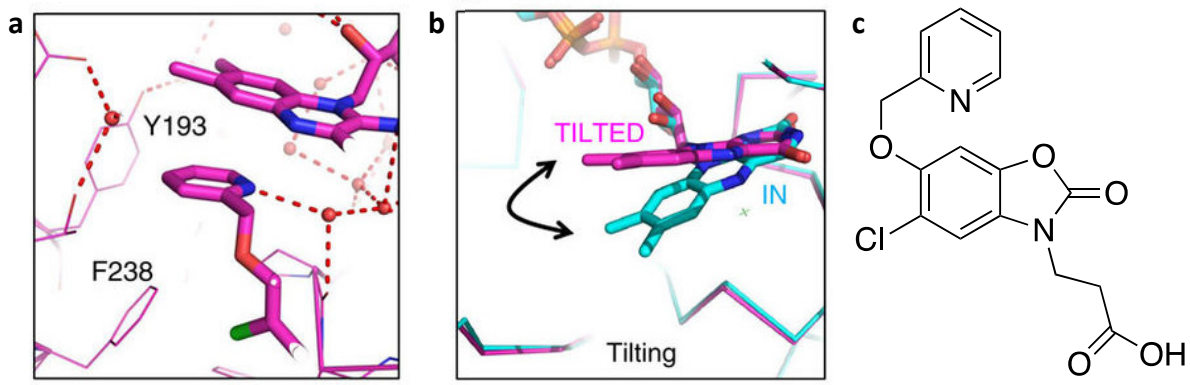

Figure 6.2 Crystal structure of dm2 bound with inhibitor GSK891. **a.** Pyridyl group of GSK891 π-stacking with isoalloxazine of FAD. **b.** Overlay of crystal structures of dm2, exhibiting the tilt observed in FAD positioning (magenta is bound with inhibitor GSK891, cyan is bound with inhibitor GSK428). **c.** Structure of GSK891. Interactions are shown as dashed lines. Image adapted from the paper by Hutchinson et al. with permission from the authors[133].

Previous examples of KMO inhibition would stimulate peroxide generation since the flavin-C4a-hydroperoxide was still formed; turnover was then interrupted by the inhibitor blocking access of the substrate. Inhibitors like GSK891, however, do not stimulate peroxide production and in fact reduce basal levels. This implies a new form of inhibition, one where the flavin-C4a-hydroperoxide is not formed. It is hypothesised that the positioning of these pendant groups in inhibitors like GSK891 blocks NADPH access to the FAD, so that FAD cannot be reduced, this is shown in Figure 6.3. This could lead to a new type of inhibition in KMO and potentially flavoproteins. Inhibitors that reduce peroxide formation are far more applicable in treating diseases, particularly neurological diseases like those associated with the KP.

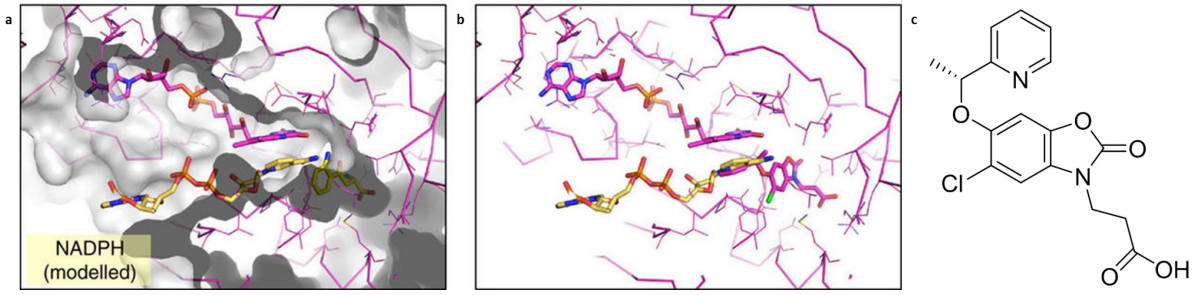

Figure 6.3 Crystal structure of dm2 bound with inhibitor GSK775. **a.** NADPH and L-Kyn (yellow) modelled into active site of the protein with GSK775 removed. **b.** Inhibitor GSK775 occupying the same space as modelled NADPH. **c.** Structure of GSK775. Image adapted from the paper by Hutchinson et al. with permission from the authors[133]. PDB 5NAE.

Sadly the NADPH shown in Figure 6.3 is only modelled. This work would benefit significantly from a structure of KMO or a type A flavoprotein aromatic hydroxylase bound with NADPH showing an

interaction between the flavin and NADPH. A confirmation of binding positioning of NADPH would allow for further development of this new form of inhibition.

References

1. Beadle, G.W., H.K. Mitchell, and J.F. Nyc, *Kynurenine as an intermediate in the formation of nicotinic acid from tryptophane by neurospora.* **Proceedings of the National Academy of Sciences of the United States of America**, 1947. 33(6): p. 155-158.
2. Ellinger, A., *The development of kynurenic acid.* **Hoppe-Seylers Zeitschrift Fur Physiologische Chemie**, 1904. 43(3/4): p. 325-337.
3. Wolf, H., *Studies on tryptophan-metabolism in man - effect of hormones and vitamin-b6 on urinary-excretion of metabolites of kynurenine pathway.* **Scandinavian Journal of Clinical & Laboratory Investigation**, 1974. 33.
4. Kang, K., K. Lee, S. Park, Y.S. Kim, and K. Back, *Enhanced production of melatonin by ectopic overexpression of human serotonin n-acetyltransferase plays a role in cold resistance in transgenic rice seedlings.* **Journal of Pineal Research**, 2010. 49(2): p. 176-182.
5. Lima, W.C., A.M. Varani, and C.F.M. Menck, *Nad biosynthesis evolution in bacteria: Lateral gene transfer of kynurenine pathway in xanthomonadales and flavobacteriales.* **Molecular Biology and Evolution**, 2009. 26(2): p. 399-406.
6. Stone, T.W., *Neuropharmacology of quinolinic and kynurenic acids.* **Pharmacological Reviews**, 1993. 45(3): p. 309-379.
7. Dai, X.C. and B.T. Zhu, *Indoleamine 2,3-dioxygenase tissue distribution and cellular localization in mice: Implications for its biological functions.* **Journal of Histochemistry & Cytochemistry**, 2010. 58(1): p. 17-28.
8. Prendergast, G.C., R. Metz, A.J. Muller, L.M.F. Merlo, and L. Mandik-Nayak, *Ido2 in immunomodulation and autoimmune disease.* **Frontiers in Immunology**, 2014. 5: p. 1-6.
9. Winn, P., T.W. Stone, M. Latimer, M.H. Hastings, and A.J.M. Clark, *A comparison of excitotoxic lesions of the basal forebrain by kainate, quinolinate, ibotenate, n-methyl-d-aspartate or quisqualate, and the effects on toxicity of 2-amino-5-phosphonovaleric acid and kynurenic acid in the rat.* **British Journal of Pharmacology**, 1991. 102(4): p. 904-908.
10. Bellac, C.L., R.S. Coimbra, S. Christen, and S.L. Leib, *Pneumococcal meningitis causes accumulation of neurotoxic kynurenine metabolites in brain regions prone to injury.* **Neurobiology of Disease**, 2006. 24(2): p. 395-402.
11. Mole, D.J., S.P. Webster, I. Uings, X.Z. Zheng, M. Binnie, K. Wilson, J.P. Hutchinson, O. Mirguet, A. Walker, B. Beaufils, N. Ancellin, L. Trottet, V. Beneton, C.G. Mowat, M. Wilkinson, P. Rowlands, C. Haslam, A. McBride, N.Z.M. Homer, J.E. Baily, M.G.F. Sharp, O.J. Garden, J. Hughes, S.E.M. Howie, D.S. Holmes, J. Liddle, and J.P. Iredale, *Kynurenine-3-monooxygenase*

inhibition prevents multiple organ failure in rodent models of acute pancreatitis. **Nature Medicine**, 2016. 22(2): p. 202-209.

12. AlberatiGiani, D., R. Buchli, P. Malherbe, C. Broger, G. Lang, C. Kohler, H.W. Lahm, and A.M. Cesura, *Isolation and expression of a cdna clone encoding human kynureninase.* **European Journal of Biochemistry**, 1996. 239(2): p. 460-468.
13. *Gene [internet]. Bethesda (md): National library of medicine (us), national center for biotechnology information; 2004 – cited 2017 sep 04. Available from: Https://www.Ncbi.Nlm.Nih.Gov/gene/883*.
14. *Gene [internet]. Bethesda (md): National library of medicine (us), national center for biotechnology information; 2004 – cited 2017 sep 04. Available from: Https://www.Ncbi.Nlm.Nih.Gov/gene/23475*.
15. Walsh, S.J. and L.M. Rau, *Autoimmune diseases: A leading cause of death among young and middle-aged women in the united states.* **American Journal of Public Health**, 2000. 90(9): p. 1463-1466.
16. Moffett, J.R. and M.A. Namboodiri, *Tryptophan and the immune response.* **Immunology and Cell Biology**, 2003. 81(4): p. 247-265.
17. Pfefferkorn, E.R., *Interferon-gamma blocks the growth of toxoplasma-gondii in human-fibroblasts by inducing the host-cells to degrade tryptophan.* **Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences**, 1984. 81(3): p. 908-912.
18. Munn, D.H., M. Zhou, J.T. Attwood, I. Bondarev, S.J. Conway, B. Marshall, C. Brown, and A.L. Mellor, *Prevention of allogeneic fetal rejection by tryptophan catabolism.* **Science**, 1998. 281(5380): p. 1191-1193.
19. Nguyen, N.T., T. Nakahama, D.H. Le, L.V. Son, H.H. Chu, and T. Kishimoto, *Aryl hydrocarbon receptor and kynurenine: Recent advances in autoimmune disease research.* **Frontiers in Immunology**, 2014. 5: p. 6.
20. Bessede, A., M. Gargaro, M.T. Pallotta, D. Matino, G. Servillo, C. Brunacci, S. Bicciato, E.M.C. Mazza, A. Macchiarulo, C. Vacca, R. Iannitti, L. Tissi, C. Volpi, M.L. Belladonna, C. Orabona, R. Bianchi, T.V. Lanz, M. Platten, M.A. Della Fazia, D. Piobbico, T. Zelante, H. Funakoshi, T. Nakamura, D. Gilot, M.S. Denison, G.J. Guillemin, J.B. DuHadaway, G.C. Prendergast, R. Metz, M. Geffard, L. Boon, M. Pirro, A. Iorio, B. Veyret, L. Romani, U. Grohmann, F. Fallarino, and P. Puccetti, *Aryl hydrocarbon receptor control of a disease tolerance defence pathway.* **Nature**, 2014. 511(7508): p. 184-190.

21. Pallotta, M.T., C. Orabona, C. Volpi, C. Vacca, M.L. Belladonna, R. Bianchi, G. Servillo, C. Brunacci, M. Calvitti, S. Bicciato, E.M.C. Mazza, L. Boon, F. Grassi, M.C. Fioretti, F. Fallarino, P. Puccetti, and U. Grohmann, *Indoleamine 2,3-dioxygenase is a signaling protein in long-term tolerance by dendritic cells.* **Nature Immunology**, 2011. 12(9): p. 870-878.

22. Munn, D.H., E. Shafizadeh, J.T. Attwood, I. Bondarev, A. Pashine, and A.L. Mellor, *Inhibition of t cell proliferation by macrophage tryptophan catabolism.* **Journal of Experimental Medicine**, 1999. 189(9): p. 1363-1372.

23. Metz, R., S. Rust, J.B. DuHadaway, M.R. Mautino, D.H. Munn, N.N. Vahanian, C.J. Link, and G.C. Prendergast, *Ido inhibits a tryptophan sufficiency signal that stimulates mtor a novel ido effector pathway targeted by d-1-methyl-tryptophan.* **Oncoimmunology**, 2012. 1(9): p. 1460-1468.

24. Terness, P., T.M. Bauer, L. Rose, C. Dufter, A. Watzlik, H. Simon, and G. Opelz, *Inhibition of allogeneic t cell proliferation by indoleamine 2,3-dioxygenase-expressing dendritic cells: Mediation of suppression by tryptophan metabolites.* **Journal of Experimental Medicine**, 2002. 196(4): p. 447-457.

25. Fallarino, I., U. Grohmann, C. Vacca, R. Bianchi, C. Orabona, A. Spreca, M.C. Fioretti, and P. Puccetti, *T cell apoptosis by tryptophan catabolism.* **Cell Death and Differentiation**, 2002. 9(10): p. 1069-1077.

26. Corthay, A., *How do regulatory t cells work?* **Scandinavian Journal of Immunology**, 2009. 70(4): p. 326-336.

27. Fallarino, F., U. Grohmann, S. You, B.C. McGrath, D.R. Cavener, C. Vacca, C. Orabona, R. Bianchi, M.L. Belladonna, C. Volpi, P. Santamaria, M.C. Fioretti, and P. Puccetti, *The combined effects of tryptophan starvation and tryptophan catabolites down-regulate t cell receptor zeta-chain and induce a regulatory phenotype in naive t cells.* **Journal of Immunology**, 2006. 176(11): p. 6752-6761.

28. Mezrich, J.D., J.H. Fechner, X.J. Zhang, B.P. Johnson, W.J. Burlingham, and C.A. Bradfield, *An interaction between kynurenine and the aryl hydrocarbon receptor can generate regulatory t cells.* **Journal of Immunology**, 2010. 185(6): p. 3190-3198.

29. Nematollahi, A., G.C. Sun, G.S. Jayawickrama, and W.B. Church, *Kynurenine aminotransferase isozyme inhibitors: A review.* **International Journal of Molecular Sciences**, 2016. 17(6): p. 22.

30. Han, Q., T. Cai, D.A. Tagle, and J.Y. Li, *Structure, expression, and function of kynurenine aminotransferases in human and rodent brains.* **Cellular and Molecular Life Sciences**, 2010. 67(3): p. 353-368.

31. Perkins, M.N. and T.W. Stone, *An iontophoretic investigation of the actions of convulsant kynurenines and their interaction with the endogenous excitant quinolinic acid.* **Brain Research**, 1982. 247(1): p. 184-187.

32. Foster, A.C., A. Vezzani, E.D. French, and R. Schwarcz, *Kynurenic acid blocks neurotoxicity and seizures induced in rats by the related brain metabolite quinolinic acid.* **Neuroscience Letters**, 1984. 48(3): p. 273-278.

33. Ganong, A.H., T.H. Lanthorn, and C.W. Cotman, *Kynurenic acid inhibits synaptic and acidic amino acid-induced responses in the rat hippocampus and spinal-cord.* **Brain Research**, 1983. 273(1): p. 170-174.

34. Vecsei, L., L. Szalardy, F. Fulop, and J. Toldi, *Kynurenines in the cns: Recent advances and new questions.* **Nat Rev Drug Discov**, 2013. 12(1): p. 64-82.

35. Zhou, Q. and M. Sheng, *Nmda receptors in nervous system diseases.* **Neuropharmacology**, 2013. 74: p. 69-75.

36. Ikonomidou, C., F. Bosch, M. Miksa, P. Bittigau, J. Vockler, K. Dikranian, T.I. Tenkova, V. Stefovska, L. Turski, and J.W. Olney, *Blockade of nmda receptors and apoptotic neurodegeneration in the developing brain.* **Science**, 1999. 283(5398): p. 70-74.

37. Kessler, M., T. Terramani, G. Lynch, and M. Baudry, *A glycine site associated with n-methyl-d-aspartic acid receptors - characterization and identification of a new class of antagonists.* **Journal of Neurochemistry**, 1989. 52(4): p. 1319-1328.

38. Nelesen, R.A., *Biological psychology: An introduction to behavioral, cognitive, and clinical neuroscience, 2nd edition.* **Biological Psychology**, 2000. 52(2): p. 185-186.

39. Prescott, C., A.M. Weeks, K.J. Staley, and K.M. Partin, *Kynurenic acid has a dual action on ampa receptor responses.* **Neuroscience Letters**, 2006. 402(1-2): p. 108-112.

40. Rozsa, E., H. Robotka, L. Vecsei, and J. Toldi, *The janus-face kynurenic acid.* **Journal of Neural Transmission**, 2008. 115(8): p. 1087-1091.

41. Levin, E.D., F.J. McClernon, and A.H. Rezvani, *Nicotinic effects on cognitive function: Behavioral characterization, pharmacological specification, and anatomic localization.* **Psychopharmacology**, 2006. 184(3-4): p. 523-539.

42. Conejero-Goldberg, C., P. Davies, and L. Ulloa, *Alpha7 nicotinic acetylcholine receptor: A link between inflammation and neurodegeneration.* **Neuroscience and Biobehavioral Reviews**, 2008. 32(4): p. 693-706.

43. Hilmas, C., E.F.R. Pereira, M. Alkondon, A. Rassoulpour, R. Schwarcz, and E.X. Albuquerque, *The brain metabolite kynurenic acid inhibits alpha 7 nicotinic receptor activity and increases*

non-alpha 7 nicotinic receptor expression: Physiopathological implications. **Journal of Neuroscience**, 2001. 21(19): p. 7463-7473.

44. Murray, I.A., A.D. Patterson, and G.H. Perdew, *Aryl hydrocarbon receptor ligands in cancer: Friend and foe.* **Nature Reviews Cancer**, 2014. 14(12): p. 801-814.

45. DiNatale, B.C., I.A. Murray, J.C. Schroeder, C.A. Flaveny, T.S. Lahoti, E.M. Laurenzana, C.J. Omiecinski, and G.H. Perdew, *Kynurenic acid is a potent endogenous aryl hydrocarbon receptor ligand that synergistically induces interleukin-6 in the presence of inflammatory signaling.* **Toxicological Sciences**, 2010. 115(1): p. 89-97.

46. Mackenzie, A.E. and G. Milligan, *The emerging pharmacology and function of gpr35 in the nervous system.* **Neuropharmacology**.

47. Wang, J.H., N. Simonavicius, X.S. Wu, G. Swaminath, J. Reagan, H. Tian, and L. Ling, *Kynurenic acid as a ligand for orphan g protein-coupled receptor gpr35.* **Journal of Biological Chemistry**, 2006. 281(31): p. 22021-22028.

48. Alkondon, M., E.F.R. Pereira, S.W. Todd, W.R. Randall, M.V. Lane, and E.X. Albuquerque, *Functional g-protein-coupled receptor 35 is expressed by neurons in the ca1 field of the hippocampus.* **Biochemical Pharmacology**, 2015. 93(4): p. 506-518.

49. Ohshiro, H., H. Tonai-Kachi, and K. Ichikawa, *Gpr35 is a functional receptor in rat dorsal root ganglion neurons.* **Biochemical and Biophysical Research Communications**, 2008. 365(2): p. 344-348.

50. Guillemin, G.J., G. Smythe, O. Takikawa, and B.J. Brew, *Expression of indoleamine 2,3-dioxygenase and production of quinolinic acid by human microglia, astrocytes, and neurons.* **Glia**, 2005. 49(1): p. 15-23.

51. Lugo-Huitron, R., P.U. Muniz, B. Pineda, J. Pedraza-Chaverri, C. Rios, and V. Perez-de la Cruz, *Quinolinic acid: An endogenous neurotoxin with multiple targets.* **Oxidative Medicine and Cellular Longevity**, 2013: p. 14.

52. Mehler, A.H., *Formation of picolinic and quinolinic acids following enzymatic oxidation of 3-hydroxyanthranilic acid.* **Journal of Biological Chemistry**, 1956. 218(1): p. 241-254.

53. Lapin, I.P., *Stimulant and convulsive effects of kynurenines injected into brain ventricles in mice.* **Journal of Neural Transmission**, 1978. 42(1): p. 37-43.

54. Stone, T.W. and M.N. Perkins, *Quinolinic acid - a potent endogenous excitant at amino-acid receptors in cns.* **European Journal of Pharmacology**, 1981. 72(4): p. 411-412.

55. Rahman, A., K. Ting, K.M. Cullen, N. Braidy, B.J. Brew, and G.J. Guillemin, *The excitotoxin quinolinic acid induces tau phosphorylation in human neurons.* **Plos One**, 2009. 4(7): p. 15.

56. Nakanishi, S., *Molecular diversity of glutamate receptors and implications for brain-function.* **Science**, 1992. 258(5082): p. 597-603.

57. Schwarcz, R. and C. Kohler, *Differential vulnerability of central neurons of the rat to quinolinic acid.* **Neuroscience Letters**, 1983. 38(1): p. 85-90.

58. Maddison, D.C. and F. Giorgini, *The kynurenine pathway and neurodegenerative disease.* **Seminars in Cell & Developmental Biology**, 2015. 40: p. 134-141.

59. Braidy, N., R. Grant, S. Adams, B.J. Brew, and G.J. Guillemin, *Mechanism for quinolinic acid cytotoxicity in human astrocytes and neurons.* **Neurotoxicity Research**, 2009. 16(1): p. 77-86.

60. Stipek, S., F. Stastny, J. Platenik, J. Crkovska, and T. Zima, *The effect of quinolinate on rat brain lipid peroxidation is dependent on iron.* **Neurochemistry International**, 1997. 30(2): p. 233-237.

61. Rodriguez-Martinez, E., A. Camacho, P.D. Maldonado, J. Pedraza-Chaverri, D. Santamaria, S. Galvan-Arzate, and A. Santamaria, *Effect of quinolinic acid on endogenous antioxidants in rat corpus striatum.* **Brain Research**, 2000. 858(2): p. 436-439.

62. Colin-Gonzalez, A.L., P.D. Maldonado, and A. Santamaria, *3-hydroxykynurenine: An intriguing molecule exerting dual actions in the central nervous system.* **Neurotoxicology**, 2013. 34: p. 189-204.

63. Notarangelo, F.M., H.Q. Wu, A. Macherone, D.R. Graham, and R. Schwarcz, *Gas chromatography/tandem mass spectrometry detection of extracellular kynurenine and related metabolites in normal and lesioned rat brain.* **Analytical Biochemistry**, 2012. 421(2): p. 573-581.

64. Christen, S., E. Peterhans, and R. Stocker, *Antioxidant activities of some tryptophan-metabolites - possible implication for inflammatory diseases.* **Proceedings of the National Academy of Sciences of the United States of America**, 1990. 87(7): p. 2506-2510.

65. Leipnitz, G., C. Schumacher, K.B. Dalcin, K. Scussiato, A. Solano, C. Funchal, C.S. Dutra-Filho, A.T.S. Wyse, C.M.D. Wannmacher, A. Latini, and M. Wajner, *In vitro evidence for an antioxidant role of 3-hydroxykynurenine and 3-hydroxyanthranilic acid in the brain.* **Neurochemistry International**, 2007. 50(1): p. 83-94.

66. Goldstein, L.E., M.C. Leopold, X.D. Huang, C.S. Atwood, A.J. Saunders, M. Hartshorn, J.T. Lim, K.Y. Faget, J.A. Muffat, R.C. Scarpa, L.T. Chylack, E.F. Bowden, R.E. Tanzi, and A.I. Bush, *3-hydroxykynurenine and 3-hydroxyanthranilic acid generate hydrogen peroxide and promote alpha-crystallin cross-linking by metal ion reduction.* **Biochemistry**, 2000. 39(24): p. 7266-7275.

67. Cavallo, F., C. De Giovanni, P. Nanni, G. Forni, and P.L. Lollini, *2011: The immune hallmarks of cancer.* **Cancer Immunology Immunotherapy**, 2011. 60(3): p. 319-326.

68. Uyttenhove, C., L. Pilotte, I. Theate, V. Stroobant, D. Colau, N. Parmentier, T. Boon, and B.J. Van den Eynde, *Evidence for a tumoral immune resistance mechanism based on tryptophan degradation by indoleamine 2,3-dioxygenase.* **Nature Medicine**, 2003. 9(10): p. 1269-1274.

69. Pilotte, L., P. Larrieu, V. Stroobant, D. Colau, E. Dolusic, R. Frederick, E. De Plaen, C. Uyttenhove, J. Wouters, B. Masereel, and B.J. Van den Eynde, *Reversal of tumoral immune resistance by inhibition of tryptophan 2,3-dioxygenase.* **Proceedings of the National Academy of Sciences of the United States of America**, 2012. 109(7): p. 2497-2502.

70. van Baren, N. and B.J. Van den Eynde, *Tryptophan-degrading enzymes in tumoral immune resistance.* **Frontiers in Immunology**, 2015. 6: p. 9.

71. Opitz, C.A., U.M. Litzenburger, F. Sahm, M. Ott, I. Tritschler, S. Trump, T. Schumacher, L. Jestaedt, D. Schrenk, M. Weller, M. Jugold, G.J. Guillemin, C.L. Miller, C. Lutz, B. Radlwimmer, I. Lehmann, A. von Deimling, W. Wick, and M. Platten, *An endogenous tumour-promoting ligand of the human aryl hydrocarbon receptor.* **Nature**, 2011. 478(7368): p. 197-203.

72. Prendergast, G.C., *Cancer why tumours eat tryptophan.* **Nature**, 2011. 478(7368): p. 192-194.

73. Frumento, G., R. Rotondo, M. Tonetti, G. Damonte, U. Benatti, and G.B. Ferrara, *Tryptophan-derived catabolites are responsible for inhibition of t and natural killer cell proliferation induced by indoleamine 2,3-dioxygenase.* **Journal of Experimental Medicine**, 2002. 196(4): p. 459-468.

74. Vogel, C.F.A., S.R. Goth, B. Dong, I.N. Pessah, and F. Matsumura, *Aryl hydrocarbon receptor signaling mediates expression of indoleamine 2,3-dioxygenase.* **Biochemical and Biophysical Research Communications**, 2008. 375(3): p. 331-335.

75. Godin-Ethier, J., L.A. Hanafi, C.A. Piccirillo, and R. Lapointe, *Indoleamine 2,3-dioxygenase expression in human cancers: Clinical and immunologic perspectives.* **Clinical Cancer Research**, 2011. 17(22): p. 6985-6991.

76. Muller, A.J., J.B. DuHadaway, P.S. Donover, E. Sutanto-Ward, and G.C. Prendergast, *Inhibition of indoleamine 2,3-dioxygenase, an immunoregulatory target of the cancer suppression gene bin1, potentiates cancer chemotherapy.* **Nature Medicine**, 2005. 11(3): p. 312-319.

77. Rohrig, U.F., S.R. Majjigapu, P. Vogel, V. Zoete, and O. Michiein, *Challenges in the discovery of indoleamine 2,3-dioxygenase 1 (ido1) inhibitors.* **Journal of Medicinal Chemistry**, 2015. 58(24): p. 9421-9437.

78. *Clinicaltrials.Gov [internet]. Bethesda (md): National library of medicine (us). 2000 feb 29 identifier nct00567931, 1-methyl-d-tryptophan in treating patients with metastatic or*

refractory solid tumors that cannot be removed by surgery; 2007 dec 4, cited 2017 sep 5; available from : Https://clinicaltrials.Gov/ct2/show/nct00567931.

79. Qian, F., J. Villella, P.K. Wallace, P. Mhawech-Fauceglia, J.D. Tario, C. Andrews, J. Matsuzaki, D. Valmori, M. Ayyoub, P.J. Frederick, A. Beck, J.Q. Liao, R. Cheney, K. Moysich, S. Lele, P. Shrikant, L.J. Old, and K. Odunsi, *Efficacy of levo-1-methyl tryptophan and dextro-1-methyl tryptophan in reversing indoleamine-2,3-dioxygenase-mediated arrest of t-cell proliferation in human epithelial ovarian cancer.* **Cancer Research**, 2009. 69(13): p. 5498-5504.

80. Löb, S., A. Königsrainer, D. Zieker, B.L.D.M. Brücher, H.-G. Rammensee, G. Opelz, and P. Terness, *Ido1 and ido2 are expressed in human tumors: Levo- but not dextro-1-methyl tryptophan inhibits tryptophan catabolism.* **Cancer Immunology, Immunotherapy: CII**, 2009. 58(1): p. 153-157.

81. Tomek, P., B.D. Palmer, J.U. Flanagan, C.W. Sun, E.L. Raven, and L.M. Ching, *Discovery and evaluation of inhibitors to the immunosuppressive enzyme indoleamine 2,3-dioxygenase 1 (ido1): Probing the active site-inhibitor interactions.* **European Journal of Medicinal Chemistry**, 2017. 126: p. 983-996.

82. Moroni, F., P. Russi, M.A. Gallomezo, G. Moneti, and R. Pellicciari, *Modulation of quinolinic and kynurenic acid content in the rat-brain - effects of endotoxins and nicotinylalanine.* **Journal of Neurochemistry**, 1991. 57(5): p. 1630-1635.

83. Chiarugi, A., R. Carpenedo, M.T. Molina, L. Mattoli, R. Pellicciari, and F. Moroni, *Comparison of the neurochemical and behavioral-effects resulting from the inhibition of kynurenine hydroxylase and/or kynureninase.* **Journal of Neurochemistry**, 1995. 65(3): p. 1176-1183.

84. Zwilling, D., S.-Y. Huang, K.V. Sathyasaikumar, F.M. Notarangelo, P. Guidetti, H.-Q. Wu, J. Lee, J. Truong, Y. Andrews-Zwilling, E.W. Hsieh, J.Y. Louie, T. Wu, K. Scearce-Levie, C. Patrick, A. Adame, F. Giorgini, S. Moussaoui, G. Laue, A. Rassoulpour, G. Flik, Y. Huang, J.M. Muchowski, E. Masliah, R. Schwarcz, and P.J. Muchowski, *Kynurenine 3-monooxygenase inhibition in blood ameliorates neurodegeneration.* **Cell**, 2011. 145(6): p. 863-874.

85. Campesan, S., E.W. Green, C. Breda, K.V. Sathyasaikumar, P.J. Muchowski, R. Schwarcz, C.P. Kyriacou, and F. Giorgini, *The kynurenine pathway modulates neurodegeneration in a drosophila model of huntington's disease.* **Current Biology**, 2011. 21(11): p. 961-966.

86. Ross, C.A. and S.J. Tabrizi, *Huntington's disease: From molecular pathogenesis to clinical treatment.* **Lancet Neurology**, 2011. 10(1): p. 83-98.

87. Beal, M.F., N.W. Kowall, D.W. Ellison, M.F. Mazurek, K.J. Swartz, and J.B. Martin, *Replication of the neurochemical characteristics of huntingtons-disease by quinolinic acid.* **Nature**, 1986. 321(6066): p. 168-171.

88. Jauch, D., E.M. Urbanska, P. Guidetti, E.D. Bird, J.P.G. Vonsattel, W.O. Whetsell, and R. Schwarcz, *Dysfunction of brain kynurenic acid metabolism in huntingtons-disease - focus on kynurenine aminotransferases.* **Journal of the Neurological Sciences**, 1995. 130(1): p. 39-47.

89. Beal, M.F., W.R. Matson, E. Storey, P. Milbury, E.A. Ryan, T. Ogawa, and E.D. Bird, *Kynurenic acid concentrations are reduced in huntingtons-disease cerebral-cortex.* **Journal of the Neurological Sciences**, 1992. 108(1): p. 80-87.

90. Guidetti, P., R.E. Luthi-Carter, S.J. Augood, and R. Schwarcz, *Neostriatal and cortical quinolinate levels are increased in early grade huntington's disease.* **Neurobiology of Disease**, 2004. 17(3): p. 455-461.

91. Guidetti, P., G.P. Bates, R.K. Graham, M.R. Hayden, B.R. Leavitt, M.E. MacDonald, E.J. Slow, V.C. Wheeler, B. Woodman, and R. Schwarcz, *Elevated brain 3-hydroxykynurenine and quinolinate levels in huntington disease mice.* **Neurobiology of Disease**, 2006. 23(1): p. 190-197.

92. Orr, A.L., S.H. Li, C.E. Wang, H. Li, J.J. Wang, J. Rong, X.S. Xu, P.G. Mastroberardino, J.T. Greenamyre, and X.J. Li, *N-terminal mutant huntingtin associates with mitochondria and impairs mitochondrial trafficking.* **Journal of Neuroscience**, 2008. 28(11): p. 2783-2792.

93. Giorgini, F., P. Guidetti, Q.V. Nguyen, S.C. Bennett, and P.J. Muchowski, *A genomic screen in yeast implicates kynurenine 3-monooxygenase as a therapeutic target for huntington disease.* **Nature Genetics**, 2005. 37(5): p. 526-531.

94. Amori, L., P. Guidetti, R. Pellicciari, Y. Kajii, and R. Schwarcz, *On the relationship between the two branches of the kynurenine pathway in the rat brain in vivo.* **Journal of Neurochemistry**, 2009. 109(2): p. 316-325.

95. Gulaj, E., K. Pawlak, B. Bien, and D. Pawlak, *Kynurenine and its metabolites in alzheimer's disease patients.* **Advances in Medical Sciences**, 2010. 55(2): p. 204-211.

96. Citron, M., *Alzheimer's disease: Strategies for disease modification.* **Nature Reviews Drug Discovery**, 2010. 9(5): p. 387-398.

97. Smith, J.R., J.F. Jamie, and G.J. Guillemin, *Kynurenine-3-monooxygenase: A review of structure, mechanism, and inhibitors.* **Drug Discovery Today**, 2016. 21(2): p. 315-324.

98. Breda, C., K.V. Sathyasaikumar, S.S. Idrissi, F.M. Notarangelo, J.G. Estranero, G.G.L. Moore, E.W. Green, C.P. Kyriacou, R. Schwarcz, and F. Giorgini, *Tryptophan-2,3-dioxygenase (tdo) inhibition ameliorates neurodegeneration by modulation of kynurenine pathway metabolites.* **Proceedings of the National Academy of Sciences of the United States of America**, 2016. 113(19): p. 5435-5440.

99. Saito, K., T.S. Nowak, S.P. Markey, and M.P. Heyes, *Mechanism of delayed increases in kynurenine pathway metabolism in damaged brain-regions following transient cerebral-ischemia.* **Journal of Neurochemistry**, 1993. 60(1): p. 180-192.

100. Moroni, F., R. Carpenedo, A. Cozzi, E. Meli, A. Chiarugi, and D.E. Pellegrini-Giampietro, *Studies on the neuroprotective action of kynurenine mono-oxygenase inhibitors in post-ischemic brain damage.* **Developments in Tryptophan and Serotonin Metabolism**, 2003. 527: p. 127-136.

101. Mole, D.J., N.V. McFerran, G. Collett, C. O'Neill, T. Diamond, O.J. Garden, L. Kylanpaa, H. Repo, and E.A. Deitch, *Tryptophan catabolites in mesenteric lymph may contribute to pancreatitis-associated organ failure.* **British Journal of Surgery**, 2008. 95(7): p. 855-867.

102. Ogawa, T., W.R. Matson, M.F. Beal, R.H. Myers, E.D. Bird, P. Milbury, and S. Saso, *Kynurenine pathway abnormalities in parkinsons-disease.* **Neurology**, 1992. 42(9): p. 1702-1706.

103. Guillemin, G.J., V. Meininger, and B.J. Brew, *Implications for the kynurenine pathway and quinolinic acid in amyotrophic lateral sclerosis.* **Neurodegenerative Diseases**, 2005. 2(3-4): p. 166-176.

104. Wonodi, I., O.C. Stine, K.V. Sathyasaikumar, R.C. Roberts, B.D. Mitchell, L.E. Hong, Y. Kajii, G.K. Thaker, and R. Schwarcz, *Downregulated kynurenine 3-monooxygenase gene expression and enzyme activity in schizophrenia and genetic association with schizophrenia endophenotypes.* **Archives of General Psychiatry**, 2011. 68(7): p. 665-674.

105. Kotake, Y. and T. Masayama, *Studies on the intermediary metabolism of tryptophans xviii-xxiv.* **Hoppe-Seylers Zeitschrift Fur Physiologische Chemie**, 1936. 243: p. 237-244.

106. Forouhar, F., J.L.R. Anderson, C.G. Mowat, S.M. Vorobiev, A. Hussain, M. Abashidze, C. Bruckmann, S.J. Thackray, J. Seetharaman, T. Tucker, R. Xiao, L.C. Ma, L. Zhao, T.B. Acton, G.T. Montelione, S.K. Chapman, and L. Tong, *Molecular insights into substrate recognition and catalysis by tryptophan 2,3-dioxygenase.* **Proceedings of the National Academy of Sciences of the United States of America**, 2007. 104(2): p. 473-478.

107. Zhang, Y., S.A. Kang, T. Mukherjee, S. Bale, B.R. Crane, T.P. Begley, and S.E. Ealick, *Crystal structure and mechanism of tryptophan 2,3-dioxygenase, a heme enzyme involved in tryptophan catabolism and in quinolinate biosynthesis.* **Biochemistry**, 2007. 46(1): p. 145-155.

108. Huang, W., Z. Gong, J. Li, and J.P. Ding, *Crystal structure of drosophila melanogaster tryptophan 2,3-dioxygenase reveals insights into substrate recognition and catalytic mechanism.* **Journal of Structural Biology**, 2013. 181(3): p. 291-299.

109. Meng, B., D. Wu, J.H. Gu, S.Y. Ouyang, W. Ding, and Z.J. Liu, *Structural and functional analyses of human tryptophan 2,3-dioxygenase.* **Proteins-Structure Function and Bioinformatics**, 2014. 82(11): p. 3210-3216.

110. Yamamoto, S. and O. Hayaishi, *Tryptophan pyrrolase of rabbit intestine - d- and l-tryptophan-cleaving enzyme or enzymes.* **Journal of Biological Chemistry**, 1967. 242(22): p. 5260-&.

111. Batabyal, D. and S.R. Yeh, *Human tryptophan dioxygenase: A comparison to indoleamine 2,3-dioxygenase.* **Journal of the American Chemical Society**, 2007. 129(50): p. 15690-15701.

112. Soliman, H., M. Mediavilla-Varela, and S. Antonia, *Indoleamine 2,3-dioxygenase is it an immune suppressor?* **Cancer Journal**, 2010. 16(4): p. 354-359.

113. Sugimoto, H., S.I. Oda, T. Otsuki, T. Hino, T. Yoshida, and Y. Shiro, *Crystal structure of human indoleamine 2,3-dioxygenase: Catalytic mechanism of o-2 incorporation by a heme-containing dioxygenase.* **Proceedings of the National Academy of Sciences of the United States of America**, 2006. 103(8): p. 2611-2616.

114. Ball, H.J., H.J. Yuasa, C.J.D. Austin, S. Weiser, and N.H. Hunt, *Indoleamine 2,3-dioxygenase-2; a new enzyme in the kynurenine pathway.* **International Journal of Biochemistry & Cell Biology**, 2009. 41(3): p. 467-471.

115. Altschul, S.F., T.L. Madden, A.A. Schaffer, J.H. Zhang, Z. Zhang, W. Miller, and D.J. Lipman, *Gapped blast and psi-blast: A new generation of protein database search programs.* **Nucleic Acids Research**, 1997. 25(17): p. 3389-3402.

116. Sievers, F., A. Wilm, D. Dineen, T.J. Gibson, K. Karplus, W.Z. Li, R. Lopez, H. McWilliam, M. Remmert, J. Soding, J.D. Thompson, and D.G. Higgins, *Fast, scalable generation of high-quality protein multiple sequence alignments using clustal omega.* **Molecular Systems Biology**, 2011. 7: p. 6.

117. Metz, R., J.B. DuHadaway, U. Kamasani, L.D. Laury-Kleintop, A.J. Muller, and G.C. Prendergast, *Novel enzyme ido2 is the preferred biochemical target of d-1mt which relieves tumor immune suppression and stimulates tumor regression.* **Proceedings of the American Association for Cancer Research Annual Meeting**, 2008. 49: p. 606.

118. Croitoru-Lamoury, J., F.M.J. Lamoury, M. Caristo, K. Suzuki, D. Walker, O. Takikawa, R. Taylor, and B.J. Brew, *Interferon-gamma regulates the proliferation and differentiation of mesenchymal stem cells via activation of indoleamine 2,3 dioxygenase (ido).* **Plos One**, 2011. 6(2): p. 13.

119. Austin, C.J.D., B.M. Mailu, G.J. Maghzal, A. Sanchez-Perez, S. Rahlfs, K. Zocher, H.J. Yuasa, J.W. Arthur, K. Becker, R. Stocker, N.H. Hunt, and H.J. Ball, *Biochemical characteristics and*

inhibitor selectivity of mouse indoleamine 2,3-dioxygenase-2. **Amino Acids**, 2010. 39(2): p. 565-578.

120. Millett, E.S., I. Efimov, J. Basran, S. Handa, C.G. Mowat, and E.L. Raven, *Heme-containing dioxygenases involved in tryptophan oxidation.* **Current Opinion in Chemical Biology**, 2012. 16(1-2): p. 60-66.

121. Hamilton, G.A., *Mechanisms of 2-electron and 4-electron oxidations catalyzed by some metalloenzymes.* **Advances in Enzymology and Related Areas of Molecular Biology**, 1969. 32: p. 55-&.

122. Chauhan, N., S.J. Thackray, S.A. Rafice, G. Eaton, M. Lee, I. Efimov, J. Basran, P.R. Jenkins, C.G. Mowat, S.K. Chapman, and E.L. Raven, *Reassessment of the reaction mechanism in the heme dioxygenases.* **Journal of the American Chemical Society**, 2009. 131(12): p. 4186-+.

123. Chung, L.W., X. Li, H. Sugimoto, Y. Shiro, and K. Morokuma, *Density functional theory study on a missing piece in understanding of heme chemistry: The reaction mechanism for indoleamine 2,3-dioxygenase and tryptophan 2,3-dioxygenase.* **Journal of the American Chemical Society**, 2008. 130(37): p. 12299-12309.

124. Chung, L.W., X. Li, H. Sugimoto, Y. Shiro, and K. Morokuma, *Oniom study on a missing piece in our understanding of heme chemistry: Bacterial tryptophan 2,3-dioxygenase with dual oxidants.* **Journal of the American Chemical Society**, 2010. 132(34): p. 11993-12005.

125. Lewis-Ballester, A., D. Batabyal, T. Egawa, C.Y. Lu, Y. Lin, M.A. Marti, L. Capece, D.A. Estrin, and S.R. Yeh, *Evidence for a ferryl intermediate in a heme-based dioxygenase.* **Proceedings of the National Academy of Sciences of the United States of America**, 2009. 106(41): p. 17371-17376.

126. Booth, E.S., J. Basran, M. Lee, S. Handa, and E.L. Raven, *Substrate oxidation by indoleamine 2,3-dioxygenase evidence for a common reaction mechanism.* **Journal of Biological Chemistry**, 2015. 290(52): p. 30924-30930.

127. Yeung, A.W.S., A.C. Terentis, N.J.C. King, and S.R. Thomas, *Role of indoleamine 2,3-dioxygenase in health and disease.* **Clinical Science**, 2015. 129(7): p. 601-672.

128. Capece, L., A. Lewis-Ballester, S.R. Yeh, D.A. Estrin, and M.A. Marti, *Complete reaction mechanism of indoleamine 2,3-dioxygenase as revealed by qm/mm simulations.* **Journal of Physical Chemistry B**, 2012. 116(4): p. 1401-1413.

129. Guillemin, G.J., D.G. Smith, G.A. Smythe, P.J. Armati, and B.J. Brew, *Expression of the kynurenine pathway enzymes in human microglia and macrophages.* **Developments in Tryptophan and Serotonin Metabolism**, 2003. 527: p. 105-112.

130. Erickson, J.B., E.M. Flanagan, S. Russo, and J.F. Reinhard, *A radiometric assay for kynurenine 3-hydroxylase based on the release of (h2o)-h-3 during hydroxylation of l- 3,5-h-3 kynurenine.* **Analytical Biochemistry**, 1992. 205(2): p. 257-262.

131. Nishimoto, Y., F. Takeuchi, and Y. Shibata, *Purification of l-kynurenine 3-hydroxylase by affinity chromatography.* **Journal of Chromatography**, 1979. 169(FEB): p. 357-364.

132. Nisimoto, Y., F. Takeuchi, and Y. Shibata, *Molecular-properties of l-kynurenine 3-hydroxylase from rat-liver mitochondria.* **Journal of Biochemistry**, 1977. 81(5): p. 1413-1425.

133. Hutchinson, J.P., P. Rowland, M.R.D. Taylor, E.M. Christodoulou, C. Haslam, C.I. Hobbs, D.S. Holmes, P. Homes, J. Liddle, D.J. Mole, I. Uings, A.L. Walker, S.P. Webster, C.G. Mowat, and C.W. Chung, *Structural and mechanistic basis of differentiated inhibitors of the acute pancreatitis target kynurenine-3-monooxygenase.* **Nature Communications**, 2017. 8: p. 12.

134. Uemura, T. and K. Hirai, *L-kynurenine 3-monooxygenase from mitochondrial outer membrane of pig liver: Purification, some properties, and monoclonal antibodies directed to the enzyme.* **Journal of Biochemistry**, 1998. 123(2): p. 253-262.

135. Hirai, K., H. Kuroyanagi, Y. Tatebayashi, Y. Hayashi, K. Hirabayashi-Takahashi, K. Saito, S. Haga, T. Uemura, and S. Izumi, *Dual role of the carboxyl-terminal region of pig liver l-kynurenine 3-monooxygenase: Mitochondrial-targeting signal and enzymatic activity*.* **Journal of Biochemistry**, 2010. 148(6): p. 639-650.

136. AlberatiGiani, D., A.M. Cesura, C. Broger, W.D. Warren, S. Rover, and P. Malherbe, *Cloning and functional expression of human kynurenine 3-monooxygenase.* **Febs Letters**, 1997. 410(2-3): p. 407-412.

137. Winkler, D., M. Beconi, L.M. Toledo-Sherman, M. Prime, A. Ebneth, C. Dominguez, and I. Munoz-Sanjuan, *Development of lc/ms/ms, high-throughput enzymatic and cellular assays for the characterization of compounds that inhibit kynurenine monooxygenase (kmo).* **Journal of Biomolecular Screening**, 2013. 18(8): p. 879-889.

138. Wilson, K., D.J. Mole, M. Binnie, N.Z.M. Homer, X. Zheng, B.A. Yard, J.P. Iredale, M. Auer, and S.P. Webster, *Bacterial expression of human kynurenine 3-monooxygenase: Solubility, activity, purification.* **Protein Expression and Purification**, 2014. 95: p. 96-103.

139. Breton, J., N. Avanzi, S. Magagnin, N. Covini, G. Magistrelli, L. Cozzi, and A. Isacchi, *Functional characterization and mechanism of action of recombinant human kynurenine 3-hydroxylase.* **European Journal of Biochemistry**, 2000. 267(4): p. 1092-1099.

140. Amaral, M., C. Levy, D.J. Heyes, P. Lafite, T.F. Outeiro, F. Giorgini, D. Leys, and N.S. Scrutton, *Structural basis of kynurenine 3-monooxygenase inhibition.* **Nature**, 2013. 496(7445): p. 382-385.

141. Crozier, K.R. and G.R. Moran, *Heterologous expression and purification of kynurenine-3-monooxygenase from pseudomonas fluorescens strain 17400.* **Protein Expression and Purification**, 2007. 51(2): p. 324-333.

142. Crozier-Reabe, K.R., R.S. Phillips, and G.R. Moran, *Kynurenine 3-monooxygenase from pseudomonas fluorescens: Substrate-like inhibitors both stimulate flavin reduction and stabilize the flavin-peroxo intermediate yet result in the production of hydrogen peroxide.* **Biochemistry**, 2008. 47(47): p. 12420-12433.

143. Wilkinson, M., *Structural dynamics and ligand binding in kynurenine-3-monooxygenase.* PhD Thesis, School of Chemistry, **University of Edinburgh**, 2013.

144. van Berkel, W.J.H., N.M. Kamerbeek, and M.W. Fraaije, *Flavoprotein monooxygenases, a diverse class of oxidative biocatalysts.* **Journal of Biotechnology**, 2006. 124(4): p. 670-689.

145. Entsch, B. and W.J.H. Vanberkel, *Flavoprotein structure and mechanism .1. Structure and mechanism of para-hydroxybenzoate hydroxylase.* **Faseb Journal**, 1995. 9(7): p. 476-483.

146. Husain, M. and V. Massey, *Kinetic studies on the reaction of para-hydroxybenzoate hydroxylase - agreement of steady-state and rapid reaction data.* **Journal of Biological Chemistry**, 1979. 254(14): p. 6657-6666.

147. Cole, L.J., B. Entsch, M. Ortiz-Maldonado, and D.P. Ballou, *Properties of p-hydroxybenzoate hydroxylase when stabilized in its open conformation.* **Biochemistry**, 2005. 44(45): p. 14807-14817.

148. Wang, J., M. Ortiz-Maldonado, B. Entsch, V. Massey, D. Ballou, and D.L. Gatti, *Protein and ligand dynamics in 4-hydroxybenzoate hydroxylase.* **Proceedings of the National Academy of Sciences of the United States of America**, 2002. 99(2): p. 608-613.

149. Gatti, D.L., B.A. Palfey, M.S. Lah, B. Entsch, V. Massey, D.P. Ballou, and M.L. Ludwig, *The mobile flavin of 4-oh benzoate hydroxylase.* **Science**, 1994. 266(5182): p. 110-114.

150. Enroth, C., H. Neujahr, G. Schneider, and Y. Lindqvist, *The crystal structure of phenol hydroxylase in complex with fad and phenol provides evidence for a concerted conformational change in the enzyme and its cofactor during catalysis.* **Structure with Folding & Design**, 1998. 6(5): p. 605-617.

151. McCulloch, K.M., T. Mukherjee, T.P. Begley, and S.E. Ealick, *Structure of the plp degradative enzyme 2-methyl-3-hydroxypyridine-5-carboxylic acid oxygenase from mesorhizobium loti maff303099 and its mechanistic implications.* **Biochemistry**, 2009. 48(19): p. 4139-4149.

152. Kobayashi, J., H. Yoshida, T. Yagi, S. Kamitori, H. Hayashi, K. Mizutani, N. Takahashi, and B. Mikami, *Role of the tyr270 residue in 2-methyl-3-hydroxypyridine-5-carboxylic acid oxygenase from mesorhizobium loti.* **J Biosci Bioeng**, 2017. 123(2): p. 154-162.

153. Haslam, C., J.P. Hutchinson, and J. Liddle, *Comparison of pharmacology of pseudomonas fluorescens and human kynurenine monooxygenase (kmo) - rapidfire mass spectrometry assay.* Internal report, GlaxoSmithKline.

154. Hutchinson, J.P., P. Rowland, M.R.D. Taylor, E.M. Christodoulou, C. Haslam, C.M. Hobbs, D. Holmes, P. Homes, J. Liddle, D.J. Mole, I. Uings, A. Walker, S.P. Webster, C.G. Mowat, and C.-w. Chung, *Structural and mechanistic basis of differentiated inhibitors of kmo, a target for acute pancreatitis.* **Nature Communications**, 2017.

155. Schreuder, H.A., J.M. Vanderlaan, W.G.J. Hol, and J. Drenth, *Crystal-structure of para-hydroxybenzoate hydroxylase complexed with its reaction-product 3,4-dihydroxybenzoate.* **Journal of Molecular Biology**, 1988. 199(4): p. 637-648.

156. Bell, H., *Characterisation of the active site of kynurenine 3-monooxygenase.* PhD Thesis, School of Chemistry, **University of Edinburgh**, 2016.

157. Dawson, R.M.C., D.C. Elliot, W.H. Elliot, and K.M. Jones, *Data for biochemical research*. 3 ed. 1989, Oxford: Oxford University Press. 580.

158. Winn, M.D., C.C. Ballard, K.D. Cowtan, E.J. Dodson, P. Emsley, P.R. Evans, R.M. Keegan, E.B. Krissinel, A.G.W. Leslie, A. McCoy, S.J. McNicholas, G.N. Murshudov, N.S. Pannu, E.A. Potterton, H.R. Powell, R.J. Read, A. Vagin, and K.S. Wilson, *Overview of the ccp4 suite and current developments.* **Acta Crystallographica Section D-Biological Crystallography**, 2011. 67: p. 235-242.

159. Battye, T.G.G., L. Kontogiannis, O. Johnson, H.R. Powell, and A.G.W. Leslie, *Imosflm: A new graphical interface for diffraction-image processing with mosflm.* **Acta Crystallographica Section D-Biological Crystallography**, 2011. 67: p. 271-281.

160. Evans, P., *Scaling and assessment of data quality.* **Acta Crystallographica Section D-Biological Crystallography**, 2006. 62: p. 72-82.

161. Winter, G., *Xia2: An expert system for macromolecular crystallography data reduction.* **Journal of Applied Crystallography**, 2010. 43: p. 186-190.

162. Vagin, A. and A. Teplyakov, *Molecular replacement with molrep.* **Acta Crystallographica Section D-Biological Crystallography**, 2010. 66: p. 22-25.

163. Emsley, P., B. Lohkamp, W.G. Scott, and K. Cowtan, *Features and development of coot.* **Acta Crystallographica Section D-Biological Crystallography**, 2010. 66: p. 486-501.

164. Murshudov, G.N., P. Skubak, A.A. Lebedev, N.S. Pannu, R.A. Steiner, R.A. Nicholls, M.D. Winn, F. Long, and A.A. Vagin, *Refmac5 for the refinement of macromolecular crystal structures.* **Acta Crystallographica Section D-Biological Crystallography**, 2011. 67: p. 355-367.

165. Lovell, S.C., I.W. Davis, W.B. Adrendall, P.I.W. de Bakker, J.M. Word, M.G. Prisant, J.S. Richardson, and D.C. Richardson, *Structure validation by c alpha geometry: Phi,psi and c beta deviation.* **Proteins-Structure Function and Genetics**, 2003. 50(3): p. 437-450.

166. *The pymol molecular graphics system.* Schrödinger, L., Version 1.8.

167. Schrodinger, L., *Https://pymol.Org/dokuwiki/doku.Php?Id=command:Align*. [Web Page] 12/09/2017.

168. Henikoff, S. and J.G. Henikoff, *Amino-acid substitution matrices from protein blocks.* **Proceedings of the National Academy of Sciences of the United States of America**, 1992. 89(22): p. 10915-10919.

169. *Ncont.* Krissinel, E., European Bioinformatics Institute, Cambridge, UK.

170. Adams, P.D., P.V. Afonine, G. Bunkoczi, V.B. Chen, I.W. Davis, N. Echols, J.J. Headd, L.W. Hung, G.J. Kapral, R.W. Grosse-Kunstleve, A.J. McCoy, N.W. Moriarty, R. Oeffner, R.J. Read, D.C. Richardson, J.S. Richardson, T.C. Terwilliger, and P.H. Zwart, *Phenix: A comprehensive python-based system for macromolecular structure solution.* **Acta Crystallographica Section D-Biological Crystallography**, 2010. 66: p. 213-221.

171. Liebschner, D., P.V. Afonine, N.W. Moriarty, B.K. Poon, O.V. Sobolev, T.C. Terwilliger, and P.D. Adams, *Polder maps: Improving omit maps by excluding bulk solvent.* **Acta Crystallographica Section D-Structural Biology**, 2017. 73: p. 148-157.

172. Oliveira, S.H.P., F.A.N. Ferraz, R.V. Honorato, J. Xavier-Neto, T.J.P. Sobreira, and P.S.L. de Oliveira, *Kvfinder: Steered identification of protein cavities as a pymol plugin.* **Bmc Bioinformatics**, 2014. 15: p. 8.

173. Rowland, P. (GlaxoSmithKline), C.G. Mowat, Personal Communication, *Regarding kmo crystallography.* 2014

174. Lesley, S.A. and I.A. Wilson, *Protein production and crystallization at the joint centerfor structural genomics.* **Journal of Structural and Functional Genomics**, 2005. 6(2): p. 71-79.

175. Page, R., S.K. Grzechnik, J.M. Canaves, G. Spraggon, A. Kreusch, P. Kuhn, R.C. Stevens, and S.A. Lesley, *Shotgun crystallization strategy for structural genomics: An optimized two-tiered crystallization screen against the thermotoga maritima proteome.* **Acta Crystallographica Section D-Biological Crystallography**, 2003. 59: p. 1028-1037.

176. Changeux, J.-P. and S. Edelstein, *Conformational selection or induced fit? 50 years of debate resolved.* **F1000 Biology Reports**, 2011. 3: p. 19.

Chapter 7: Appendices

7.1. Appendix 1 Crystallographic Data

7.1.1. dm2 *Pf*KMO

	Substrate-bound	Substrate-free
Beamline	I04-1	I02
X-ray wavelength (Å)	0.9282	0.9795
Space group	P1 2_1 1	P1 2_1 1
Cell dimensions *a, b, c* (Å) α, β, γ (°)	 69.93, 52.62, 137.99 90.00, 104.07, 90.00	 69.74, 51.89, 136.29 90.00, 103.69, 90.00
Resolution (Å)	67.83-1.50 (1.58-1.50)	67.76-1.79
$CC_{1/2}$	0.998 (0.495)	0.999 (0.893)
R_{merge}	0.062 (0.614)	0.047 (0.300)
Mean (I)/sd(I)	8.1 (1.7)	14.3 (4.1)
Completeness (%)	99.7 (99.6)	99.6 (99.8)
Multiplicity	3.1 (3.0)	3.4 (3.4)
R_{factor}	0.1797 (0.33)	0.1692 (0.21)
R_{free}	0.2142 (0.36)	0.2080 (0.26)
No. reflections	155570 (22560)	89963 (13032)
rmsd bond lengths (Å)	0.0228	0.0192
rmsd bond angles (°)	2.1068	1.9531
Ramachandran outliers (%)	0.81	1.40
Mean B-factors ($Å^2$) :		
Protein atoms	27.46	29.78
Ligand atoms	18.20	16.83
Waters	39.75	38.45

7.1.2. R84K *Pf*KMO

	Substrate-bound	Substrate-free
Beamline	I04-1	I03
X-ray wavelength (Å)	0.9282	0.9763
Space group	P1 2_1 1	P1 2_1 1
Cell dimensions *a, b, c* (Å) α, β, γ (°)	 70.17, 52.38, 137.64 90.00, 103.56, 90.00	 70.17, 53.16, 135.59 90.00, 104.45, 90.00
Resolution (Å)	68.22-1.70 (1.79-1.70)	67.95-1.80 (1.90-1.80)
$CC_{1/2}$	0.999 (0.486)	0.986 (0.852)
R_{merge}	0.046 (0.367)	0.113 (0.388)
Mean (I)/sd(I)	10.9 (2.3)	4.6 (1.8)
Completeness (%)	94.5 (90.0)	98.3 (98.1)
Multiplicity	2.9 (2.3)	2.7 (2.6)
R_{factor}	0.1774 (0.33)	0.1879 (0.28)
R_{free}	0.2140 (0.33)	0.2283 (0.32)
No. reflections	101305 (13991)	88433 (12831)
rmsd bond lengths (Å)	0.0200	0.0186
rmsd bond angles (°)	1.9937	1.9520
Ramachandran outliers (%)	1.25	0.35
Mean B-factors ($Å^2$) :		
Protein atoms	30.39	31.84
Ligand atoms	21.45	15.65
Waters	38.94	38.84

7.1.3. Y404F *Pf*KMO

	Substrate-bound	Substrate-free
Beamline	I04-1	I03
X-ray wavelength (Å)	0.9282	0.9763
Space group	P1 2_1 1	P1 2_1 1
Cell dimensions *a, b, c* (Å) α, β, γ (°)	 70.48, 52.99, 137.17 90.00, 104.60, 90.00	 70.24, 52.30, 136.60 90.00, 104.83, 90.00
Resolution (Å)	68.20-1.80 (1.90-1.80)	67.90-1.65 (1.74-1.65)
$CC_{1/2}$	0.997 (0.662)	0.997 (0.800)
R_{merge}	0.069 (0.449)	0.070 (0.441)
Mean (I)/sd(I)	8.4 (2.0)	8.7 (2.4)
Completeness (%)	96.0 (96.0)	99.8 (99.8)
Multiplicity	3.1 (2.5)	3.7 (3.5)
R_{factor}	0.1662 (0.30)	0.1748 (0.25)
R_{free}	0.2034 (0.33)	0.2099 (0.28)
No. reflections	87357 (12690)	115511 (16759)
rmsd bond lengths (Å)	0.0196	0.0201
rmsd bond angles (°)	1.9149	2.0680
Ramachandran outliers (%)	0.46	1.50
Mean B-factors ($Å^2$) :		
Protein atoms	31.79	33.15
Ligand atoms	27.98	21.12
Waters	42.66	40.21

7.1.4. H320F *Pf*KMO

	Substrate-bound	Substrate-free
Beamline	I03	I03
X-ray wavelength (Å)	0.9700	0.9700
Space group	P1 2_1 1	P1 2_1 1
Cell dimensions *a, b, c* (Å) α, β, γ (°)	 69.09, 51.60, 136.28 90.00, 103.96, 90.00	 69.44, 51.89, 135.95 90.00, 103.94, 90.00
Resolution (Å)	34.54-1.89	34.72-1.90 (2.01-1.90)
$CC_{1/2}$	0.999 (0.985)	0.995 (0.919)
R_{merge}	0.049 (0.132)	0.079 (0.303)
Mean (I)/sd(I)	21.3 (9.4)	9.1 (3.4)
Completeness (%)	99.8 (98.9)	99.7 (98.8)
Multiplicity	6.0 (5.2)	4.3 (3.8)
R_{factor}	0.1550 (0.17)	0.1796 (0.23)
R_{free}	0.2005 (0.24)	0.2272 (0.30)
No. reflections	74535 (10712)	74011 (10657)
rmsd bond lengths (Å)	0.0188	0.0184
rmsd bond angles (°)	1.9466	1.9007
Ramachandran outliers (%)	0.24	0.70
Mean B-factors ($Å^2$) :		
Protein atoms	27.06	38.51
Ligand atoms	23.71	24.33
Waters	37.20	40.85

7.1.5. R386K *Pf*KMO

	Substrate-exposed	Substrate-free
Beamline	I03	I03
X-ray wavelength (Å)	0.9763	0.9763
Space group	P1 2_1 1	P1 2_1 1
Cell dimensions *a, b, c* (Å) α, β, γ (°)	 66.12, 48.97, 129.36 90.00, 103.69, 90.00	 70.16, 51.83, 136.72 90.00 103.62 90.00
Resolution (Å)	29.02-1.65 (1.69-1.65)	68.19-1.89 (2.00-1.89)
$CC_{1/2}$	0.998 (0.694)	0.996 (0.840)
R_{merge}	0.054 (0.482)	0.066 (0.323)
Mean (I)/sd(I)	11.3 (2.1)	8.9 (2.9)
Completeness (%)	99.5 (98.3)	99.4 (98.6)
Multiplicity	3.3	2.9 (2.6)
R_{factor}	0.1720 (0.25)	0.1770 (0.25)
R_{free}	0.2095 (0.28)	0.2240 (0.30)
No. reflections	96781 (7014)	76115 (10975)
rmsd bond lengths (Å)	0.0396	0.0191
rmsd bond angles (°)	2.6124	1.9490
Ramachandran outliers (%)	0.47	0.91
Mean B-factors ($Å^2$) :		
Protein atoms	26.72	34.10
Ligand atoms	19.01	21.98
Waters	34.86	39.09

7.1.6. E372T *Pf*KMO

	Substrate-exposed	Substrate-free
Beamline	I02	I02
X-ray wavelength (Å)	0.9795 Å	0.9795 Å
Space group	P1 2_1 1	P1 2_1 1
Cell dimensions *a, b, c* (Å) α, β, γ (°)	 69.81, 51.21, 136.41 90.00, 104.13, 90.00	 69.50, 50.80, 136.03 90.00, 104.46, 90.00
Resolution (Å)	67.69-1.79 (1.89-1.79)	67.30-1.88 (1.99-1.88)
$CC_{1/2}$	0.998 (0.827)	0.994 (0.852)
R_{merge}	0.047 (0.437)	0.085 (0.481)
Mean (I)/sd(I)	11.0 (2.6)	6.3 (2.0)
Completeness (%)	99.5 (99.7)	99.1 (98.4)
Multiplicity	3.1 (3.1)	3.4 (3.4)
R_{factor}	0.1856 (0.28)	0.1968 (0.31)
R_{free}	0.2310 (0.28)	0.2423 (0.32)
No. reflections	88116 (12793)	73816 (10620)
rmsd bond lengths (Å)	0.0179	0.0175
rmsd bond angles (°)	1.8655	1.7968
Ramachandran outliers (%)	1.16	0.90
Mean B-factors ($Å^2$) :		
Protein atoms	44.01	45.26
Ligand atoms	27.51	28.18
Waters	45.88	46.03